2017

Search for Type III Seesaw Heavy Leptons in proton-proton collisions with the ATLAS detector at the LHC

Daniel Pluth
Iowa State University

Search for Type III Seesaw Heavy Leptons in proton-proton collisions with the ATLAS detector at the LHC

by

Daniel Jeffrey Pluth

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: High Energy Physics

Program of Study Committee:
Soeren Prell, Major Professor
Chunhui Chen
Mani Mina
Marzia Rosati
Kerry Whisnant

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2017

DEDICATION

I would like to dedicate this thesis to my wife Thaís without whose support I would not have been able to complete this work.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my gratitude to those who helped me with various aspects of my research and the writing of this thesis.

Firstly, I'd like to thank Professor Soeren Prell for his patience, insight, and attention to detail. His guidance helped me make this the most thorough and complete work of my life thus far. I would also like to thank the ISU ATLAS group for their contributions, assistance, and friendship. Thank you to my friends and family, for commiserating with my challenges and providing unending support. Finally, a very special thank you to my wife. She has been a great fount of motivation and guidance which I can certainly lack on my own.

ABSTRACT

A search for pair-production of heavy leptons (Σ^0, $\Sigma^{\pm}$) predicted by the Type III seesaw mechanism to explain the origin of small neutrino masses is presented. The decay channels for the heavy leptons $\Sigma^0 \to W^{\pm}e^{\mp}$ and $\Sigma^{\pm} \to W^{\pm}\nu_e$ are considered. Multivariate techniques versus a kinematic selection method are compared. The study is based on proton-proton collisions recorded at the ATLAS detector at the Large Hadron Collider at CERN during 2015 with a center-of-mass energy of 13 TeV. The data used corresponds to an integrated luminosity of 3.2 fb^{-1} and no evidence of heavy lepton pair-production was observed.

CHAPTER 1. INTRODUCTION

One of the basic questions that humans have been trying to answer since the beginning of history is "What comprises the world around us?". Perhaps an obvious way to answer this question is to break apart the world into the smallest possible pieces and to identify them. In the 5^{th} century BC, the Greek philosopher Democritus proposed that all matter was composed of the same 'atoms', indivisible particles comprising all matter [1].

Atoms may answer the structural aspect of that question, but the other side is about substance, what makes some matter distinct from other matter? This was dealt with the concept of 'elements'. The idea being that all matter could be composed of combinations of the most basic elements. The modern elements have been cataloged since Dmitri Mendeleev proposed the periodic table in 1869 [2]. In the 1800's, John Dalton united the concepts of atomism and elements by proposing that each of the elements are materials made from atoms of the same type [3].

In 1917, Ernest Rutherford, proved that the nucleus of the hydrogen atom is present in other nuclei [4]. This means that the atoms of all elements carry a common constituent, now called the proton. The significance of this discovery was that an indivisible atom is divisible. Soon afterward a neutral nucleus particle, the neutron, was discovered by James Chadwick in 1932 [5]. Since then, even these particles have been found to have constituents, called quarks [6].

In 1930, a hitherto unknown particle was proposed by Wolfgang Pauli in order to explain a phenomenon observed when a neutron decays to a proton and an electron (beta decay) [7]. In such two-body decays, where the neutron is initially at rest, the energies of the electron and the proton have unique values. However, when measured, it was discovered that beta decays produce electrons with a continuous spectrum of energy. This discovery strongly

suggested the existence of a third, neutral particle in beta decay, termed the neutrino by Enrico Fermi. In addition to the neutrino produced in neutron beta decay, two more neutrinos have been discovered since. However, despite the first neutrino discovery more than 80 years ago [8], neutrinos are still not fully understood. For example, the current accepted theoretical framework known as the Standard Model had assumed that neutrinos were massless, however experiments in the last few decades have proven that they are not [9].

My work focuses on the origin of neutrino masses. There are many models [10–15] which could account for the neutrino masses. One such model, the Seesaw Mechanism [16], predicts the existence of a heavy sibling to the neutrinos in the Standard Model which would act as a counter-balance to allow the Standard Model neutrinos to acquire their very small mass. Until recently, this model was not testable due to the high energies required to create the sibling neutrino. However, with the Large Hadron Collider at CERN, the lowest energy ranges of this theory are beginning to be examined.

CHAPTER 2. THE STANDARD MODEL

This chapter summarizes the current state of the Standard Model. Particles are discussed first, followed by the forces that act between the particles. There will be a focus on neutrinos, and their unique properties. At last, the chapter discusses some current problems with the Standard Model.

2.1 Introduction

The Standard Model (SM) describes all known elementary particles and the interactions between them. The most elementary particles of the Standard Model are tabulated in Figure 2.1. The atoms which make up the matter around us are formed from electrons, protons and neutrons. These protons and neutrons are made up of the lowest mass quarks; the up quark and the down quark. Four of the particles shown in Figure 2.1 are known as force carrying particles: the W and Z bosons, the photon, and the gluon. Additionally the graviton would belong to these force carrying particles, although it should be noted that the graviton is not a part of the Standard Model and that there is no accepted quantum mechanical explanation for gravity.

Each charged particle and some force carrying particles, namely the W boson, has an associated anti-particle with the same mass but opposite charge. The Higgs boson sits between the electromagnetic and weak forces to represent the fact that the Higgs field breaks the electroweak symmetry resulting in the electromagnetic and weak forces. The Higgs field is described in more detail in Section 2.5.

Additionally, the figure shows which forces interact with which particles. All other particles in the SM are comprised of one or more of these particles or their anti-particles. The SM has been tested extensively since its inception and has been very effective predictor

of elementary interactions. However it is not complete, there are still many phenomena that need further investigation and study. A few of those blind spots are discussed in Section 2.7, which covers physics beyond the Standard Model.

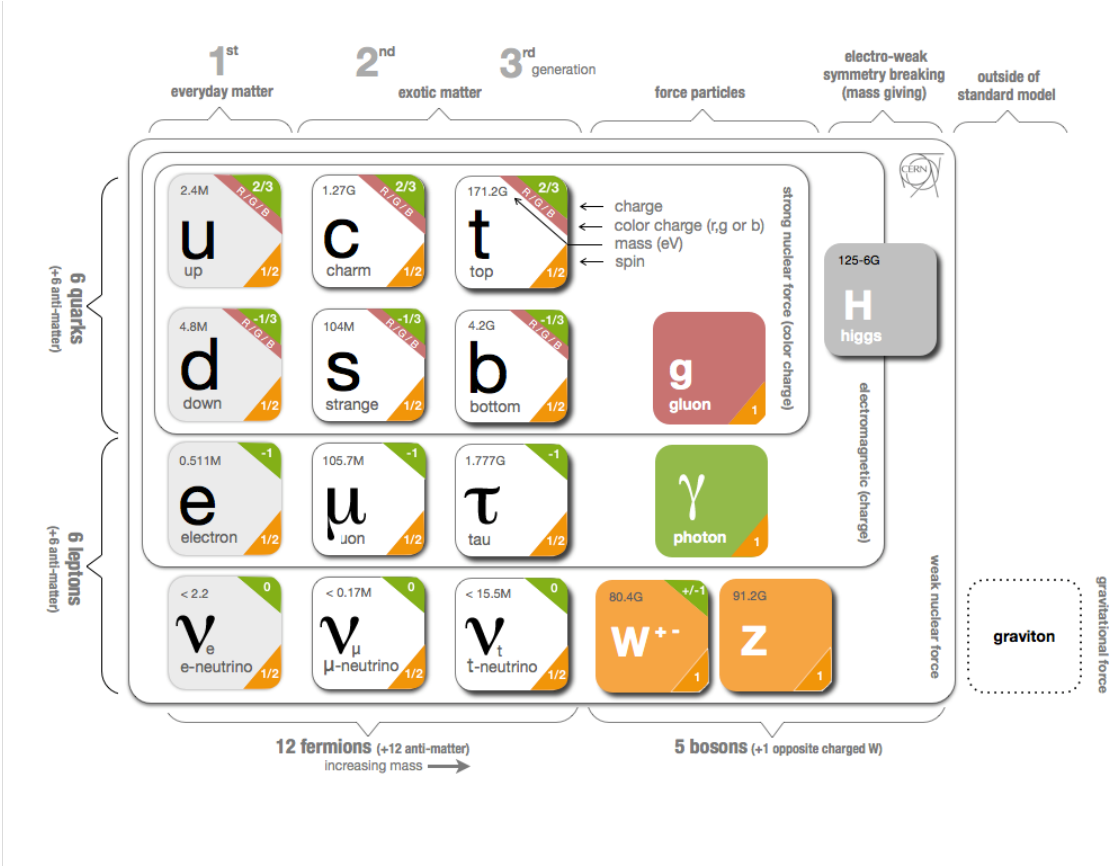


Figure 2.1: Diagram of the elementary particles in the Standard Model [17].

2.2 Quarks

Protons and neutrons are the most common examples of baryons, which are particles made up of three quarks or three anti-quarks. Mesons are particles made up of a quark and an anti-quark. Mesons and baryons, as well as tetraquarks and pentaquarks, are categorized together as hadrons. Quarks are not free particles, and thus are always observed as part of a hadron. This confinement is discussed further in Section 2.4.2.

All quarks are fermions, which is defined as any particle which has a half integer spin. Spin is the intrinsic angular momentum of a particle. There are six flavors of quarks, 3 up-type and 3 down-type. The up-type have a charge of $+2/3$ e and the down-type have a charge of $-1/3$ e, where e is the elementary charge, approximately 1.6×10^{-19} coulombs. Thus two up quarks and a down quark form a proton with a charge of $+1$ e, and an up quark with two down quarks form a neutron with no charge. In addition, each of these quarks also has an anti-particle making for a total of 12 quarks.

All quarks interact via the electromagnetic, weak, and strong interactions. Quarks, along with leptons are grouped into generations which contain: an up-type quark, a down-type quark, a charged lepton, and a neutrino. Each subsequent generation contains a corresponding set of heavier mass particles with the same charges as the previous generation, as is shown in Figure 2.1. In each generation the interactions are the same but the mass and flavor quantum numbers differ.

2.3 Leptons

Electrons, muons, and τ particles are the three charged leptons in the Standard Model. Leptons have $\pm\frac{1}{2}\hbar$ spin and therefore are fermions. Each of these leptons carries an electric charge of $-1e$. The electron is the most familiar of the leptons, the muon is a heavier flavor, and the τ is heavier still. The electron has a mass of 510 keV, the muon 105 MeV and the τ 1.7 GeV. The electron is the lightest lepton, and therefore is stable, however the muon has a lifetime of 2.2×10^{-6} seconds, and the τ has a lifetime of 290.3×10^{-15} seconds [18]. The τ is sufficiently massive that it is able to decay semi-leptonically to final states including mesons via a W boson, in fact, 65% of the time τ's decay to pions. The remaining decay fraction is fully leptonic via a W boson to a charged lepton and two neutrinos. W bosons

will be discussed further in Sections 2.4.3. The W in this decay is virtual, also called off mass shell or off-shell[†].

The remaining 35% is split between leptonic final states. The muon itself is only energetically permitted to decay to an electron, an anti-electron neutrino, and a muon neutrino [18].

For each charged lepton, there is an associated neutrino, a neutral and extremely lightweight particle. All leptons interact through the weak force. The charged leptons also interact through the electromagnetic force. That makes for a total of 12 leptons in the Standard Model, since each of these particles also has an oppositely charged anti-particle.

2.3.1 Neutrinos

Neutrinos are the neutral leptons in the Standard Model. Neutrinos only interact through the weak force, and because of this the neutrino scattering cross-section is extremely small. Cross-section in particle physics generally means the likelihood of interaction between two particles. The initial detection of neutrinos found a cross-section of only 6.3×10^{-48} m^2 [19]. The mean free path for such neutrinos through lead is calculated in Equation 2.2. The cross-section (σ) of interaction is heavily dependent on the energy of the neutrino. N is the number density, the density of lead is given by ρ, and u is the mass of a lead atom.

$$\lambda = \frac{1}{\sigma N} = \frac{u}{\sigma \rho} = \frac{3.44 * 10^{-25}\,\mathrm{kg}}{6.3 * 10^{-48}\,\mathrm{m}^2 \quad 11.3 * 10^{4}\,\mathrm{kg/m}^3} = 4.8 * 10^{17}\mathrm{m} \tag{2.2}$$

The result being, if a sample of neutrinos were passed through a block of lead, roughly a light-year in length, only 50% would interact with the lead over such a distance. Be-

[†]Equation 2.1 shows the energy-momentum relation which real particles obey. However, when a particle is off-shell, the left-hand side of the equation does not equal the mass. This is permitted, however the amplitude of the interaction diminishes the further off-shell the particle is.

$$E^2 - p^2c^2 = m^2c^4 \tag{2.1}$$

cause of their small scattering cross-section, the neutrino some of the most difficult stable particles to detect. Direct detection of an expected neutrino in a single interaction is essentially impossible. As a consequence neutrino experiments rely on indirect measurements or statistical studies.

2.4 Forces

There are four known forces in the Standard Model: the gravitational force, the electromagnetic force, the weak force, and the strong force. Some effects of the gravitational and electromagnetic forces are obvious, from the trajectory of a thrown ball, or when playing with magnets. However the strong and weak forces are not noticeable in daily life. Each of the forces have an associated force carrier particle.

2.4.1 Electromagnetic Force

The photon is the force carrier of the electromagnetic force. It has no mass in the Standard Model, and it travels at the speed of light. It has a spin of $\pm\hbar$ and has no electric charge. Charge is a measure of the strength of the particle's reaction with its associated force. Since the photon is massless, it can not decay to another particle. The photon interacts with all charged particles via the electromagnetic force.

2.4.2 Strong Force

The nucleus of an atom consists of positively charged and neutral particles. Without the strong force, the nucleus would tear itself apart due to the electro-static repulsion between the protons. The force of gravity is far weaker than the electromagnetic at such small lengths, so it is not sufficient to bind the atom, it is only because of the strong force that the atom is bound together. The strong force is an force which governs quark and gluon interactions, which will be discussed later. The effect of quark and gluon interactions is largely confined within the protons and neutrons or 'nucleons'.

Gluons are the carriers of the strong force. They have an intrinsic property known as color or color charge, which is somewhat an analogue to the electric charge of the electromagnetic force. The three colors are chosen to be red, blue, and green; such that the combination of the three is 'white' or colorless. The theory that governs these interactions is called quantum chromodynamics or QCD.

Quarks are not free particles, all quarks are confined to hadrons. The term for the non-isolation of quarks is color confinement. Confinement implies that a single quark can not be isolated, in fact all physical objects are believed to be color neutral. Both gluons and quarks are color-charged. When one tries to isolate a quark, it becomes more energetically favorable to spontaneously generate a quark anti-quark pair than to pull apart the existing quarks. These newly created quarks will bind with the separating quarks and form new mesons or baryons. The process of spontaneous generation of secondary particles from quarks or gluons is called hadronization.

A meson might be composed of an up-type quark and an anti up-type quark. These quarks must have complementary color, for example: one being red, and the other anti-red. A baryon becomes color neutral with the three color states, blue, green, and red, or anti-blue, anti-green, and anti-red.

There is one exception to color confinement, namely the top quark. The top quark's lifetime is a mere 5×10^{-25} seconds [18]. This time frame is shorter than the typical timescale of the strong interaction, which governs hadronization. Thus, the top quark offers a unique ability to study an isolated quark as it decays before it can hadronize.

2.4.3 Weak Force

The weak force governs the hydrogen to deuterium fusion in stars, and β decay. β^- decay is where a neutron turns into a proton and releases an electron and an electron anti$-\nu$. The weak force does not respect several properties such as parity and charge-parity which, due to their conservation in strong and electromagnetic interactions, were believed to be universal.

The weak force combines with the electromagnetic force at large energies to form the electroweak force, with four associated massless bosons. Bosons are zero or integral spin particles and all known force carriers are bosons. The Higgs mechanism, discussed further in Section 2.5, causes a spontaneous symmetry breaking which separates the electroweak into the weak and electromagnetic forces. This spontaneous symmetry breaking is responsible for the mass of the bosons associated with the weak force, namely the Z, W^-, and W^+. The photon of the electromagnetic force does not acquire a mass in this mechanism. The preceding '$^-$' signifies that the particle has the equivalent charge of the electron, and the '$^+$' is the negative of the electron's charge.

The W^+, W^-, and Z bosons are the force carriers of the weak force. The W and Z bosons were predicted around 1968 to create a unified theory of electromagnetism and the weak force [20–22]. The W boson governs β nuclear decays, but the Z boson had never before been observed. In 1983 on-shell W bosons were produced and observed at CERN [23] with the Super Proton Synchrotron and just a few months later the Z boson was also discovered [24, 25].

W and Z bosons are created in roughly one in every million collisions at the LHC, and can decay leptonically or hadronically. The Z boson, being neutral will decay to: leptons and anti-leptons, neutrinos and anti-neutrinos, or quark and anti-quarks. The W bosons decay similarly, but with a positive or negative charge total, so for Ws decaying leptonically this means decaying to a lepton and a neutrino. Additionally, according to the SM, the W and Z bosons always conserve lepton flavor, meaning $Z \to e^- e^+$, and never $Z \to e^- \mu^+$.

The W boson also interacts with quarks, shown in Figure 2.2, where the likelihood of a given interaction is governed by the Cabibbo–Kobayashi–Maskawa (CKM) matrix [18]. This matrix describes the strength of flavor changing decays. The matrix is mostly diagonal, which mean same generation quark interactions are favored, for example, the top quark decays almost exclusively to a bottom quark and a W boson. There is However, quark

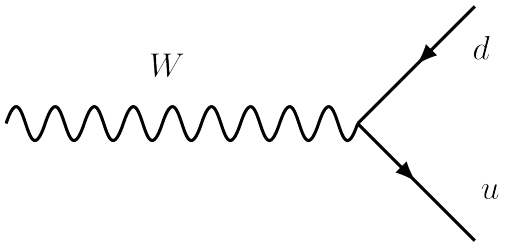


Figure 2.2: Feynman diagram showing a hadronic decay of the W boson. A Feynman diagram is a toy diagram representing the linear integrations which describe the amplitude of a particular interaction. Typically the x-axis represents time, and the y-axis is spacial.

generation mixing is allowed via the non-zero off-diagonal matrix elements, for example $W^+ \to c\bar{d}$.

$$\begin{bmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{bmatrix} = \begin{bmatrix} 0.97417 \pm 0.00021 & 0.2248 \pm 0.0006 & (4.09 \pm 0.39) \times 10^{-3} \\ 0.220 \pm 0.005 & 0.995 \pm 0.016 & (40.5 \pm 1.5) \times 10^{-3} \\ (8.2 \pm 0.6) \times 10^{-3} & (40.0 \pm 2.7) \times 10^{-3} & 1.009 \pm 0.031 \end{bmatrix}$$

2.5 Higgs Boson

The Higgs mechanism is responsible for the spontaneous symmetry breaking mentioned in Section 2.4, without which all bosons would be massless. The strength of couplings to the Higgs field governs how much mass is acquired by quarks, leptons and the W and Z bosons. The Higgs coupling governs the Higgs interaction with bosons, while the couplings to quarks and leptons are known as Yukawa couplings.

The Higgs boson is able to decay to two W bosons or two Z bosons, as these particles can be produced off-shell. In decays to dibosons (WW or ZZ), one of the bosons is always off-shell, which, as mentioned earlier, diminishes the likelihood of the decay. The likelihoods of a particle's decays can also be represented as branching fractions. The branching fractions of all a particle's decays sum to unity.

The Higgs boson does not directly couple to photons, however it can still decay into a pair of photons via virtual particles as shown in Figure 2.3. This Feynman diagram has a loop of virtual particles. In this case, several particles are possible within the loop, namely

all the quarks are possible. Since the particles in the loop are virtual, the top quark can be produced in the loop, even though it is more massive than the Higgs. Each of the possibilities contribute to the total amplitude of the interaction. The cross-section can be derived from the amplitude. Any significant discrepancy in cross-section from Standard Model predictions could imply a fourth generation of quarks or other exotic particles.

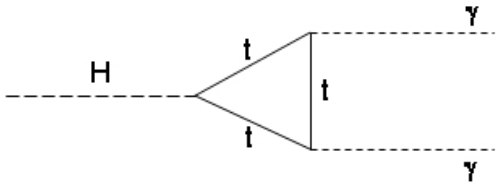


Figure 2.3: Higgs decay to two photons via a virtual top quark within a loop.

The Higgs particle was first predicted in 1964 [26]. The minimal version of the Higgs mechanism only requires one Higgs boson, however there are many theories [27–29] which include an expanded Higgs sector, for example: Higgs doublets, composite Higgs, or charged Higgses. On July 4th, 2012, a particle consistent with the neutral-charged Higgs was discovered at the Large Hadron Collider (LHC), with a mass of 125 gigaelectronvolts (GeV) [30,31]. The spin and branching fractions also seem consistent with the Standard Model. Research continues in order to understand other properties of this particle and to ascertain if there are additional charged or neutral Higgs bosons.

2.6 Neutrino Masses

The Standard Model does not provide for neutrino masses. The mass term in the Yukawa interaction with the Higgs cancels out, leaving a massive corresponding left-handed charged fermion, but a massless neutral fermion. This seems consistent with observation since neutrinos have always been observed to be left-handed. Handedness or helicity, is the projection of spin onto the momentum vector. All observed neutrinos have negative

helicity, also known as left-handedness. The first experiment to prove this was the Goldhaber experiment diagrammed in Figure 2.4 [32].

The Goldhaber experiment takes advantage of convenient energy levels in ^{152}Sm and the β-capture of ^{152}Eu. The ^{152}Eu β-captures to form ^{152}Sm*, which releases a photon. The $^{152}Sm_2O_3$ material in Figure 2.4 has a resonance at the energy of this produced photon, and can be re-emitted and detected by the scintillator. The helicity of the photon can be selected by flipping the magnetic analyzer, as the spin of the atoms in the analyzer would align with the magnetic field. A photon scattered through an oppositely polarized analyzer will lose energy and no longer be able to excite the material. The helicity of the photon gives the helicity of the neutrino since angular momentum is conserved. The Goldhaber experiment found that the ν_e helicity is -1.0 ± 0.3.

If neutrinos are massless and therefore travel at the speed of light, then no matter what reference frame is chosen, they will always have the same helicity.

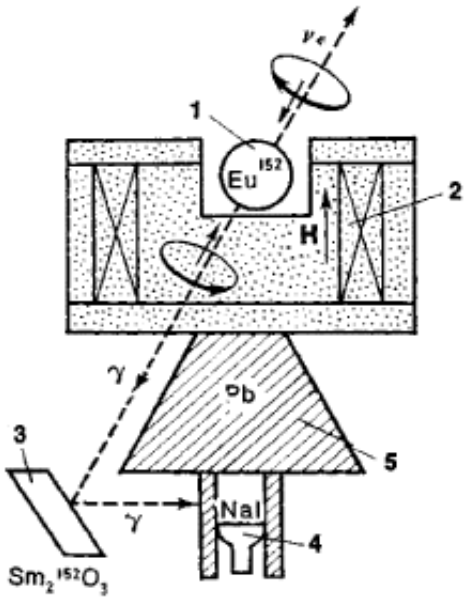


Figure 2.4: Diagram of experiment to measure helicity of neutrinos: (1) radioactive preparation ^{152}Eu, (2) magnetic analyzer (magnetized iron), (3) ^{152}Sm nuclei, (4) Scintillation counter, (5) lead shield.

2.6.1 Neutrino Oscillations

Neutrino oscillation was proposed by Bruno Pontecorvo in 1957 [33], and first observed in the 1960's when analyzing the flux of solar neutrinos. The total expected flux from the sun can be estimated using the standard solar model [34]. In the model, the sun is treated as a sphere of gas in hydrostatic equilibrium where the outward pressure created by the heat of the star is balanced by the gravitational force. It is also assumed that the sun was initially homogeneous, and its original composition can be estimated using spectroscopy and other methods [35]. The model is corrected using the current measured luminosity, radius, and composition of the sun.

According to the model and the known fusion reactions which sustain the sun, the majority of solar neutrinos are produced when two protons fuse to diproton. The diproton then β decays to deuterium, releasing a positron and an electron neutrino (ν_e) This reaction is responsible for the production of 86% of solar neutrinos. However, the neutrinos produced in this reaction have $\leq$ 0.42 MeV of energy, which was too little for early detectors to detect [36].

Equation 2.3, shows the solar reactions which produced the solar neutrinos which were originally measured. The expected solar neutrino flux was calculated based on these reactions.

$$\begin{aligned} {}^{3}\mathrm{He} + p &\rightarrow {}^{4}\mathrm{He} + e^{+} + \nu_e \\ {}^{7}\mathrm{Be} + e^{-} &\rightarrow {}^{7}\mathrm{Li} + \nu_e \\ {}^{8}\mathrm{B} &\rightarrow {}^{8}\mathrm{Be}^{*} + e^{+}\nu_e \end{aligned} \tag{2.3}$$

However, there was a measured deficit in the total solar neutrino flux compared to the predictions of the Standard Model. Depending on the reaction, only a third to a half of the expected flux was measured [37].

An explanation for this deficit is oscillation [9]. The idea is that there are three neutrino flavor eigenstates and 3 neutrino mass eigenstates. The mixing ratios between the states are expressed via the Pontecorvo−Maki−Nakagawa−Sakata (PMNS) matrix in Equation 2.4. A neutrino in a flavor eigenstate, for example the electron neutrino, is a superposition of the three mass eigenstates as seen in the wavefunction in Equation 2.5.

$$\begin{bmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{bmatrix} = \begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{bmatrix} \tag{2.4}$$

$$|\psi(t)\rangle = U_{e1}\,|\nu_1\rangle\, e^{-ip_1 x} + U_{e2}\,|\nu_2\rangle\, e^{-ip_2 x} + U_{e3}\,|\nu_3\rangle\, e^{-ip_3 x} \tag{2.5}$$

The p_i are the four-momenta of the mass eigenstates [38]. Oscillation probabilities can also be calculated, usually by re-arranging and expressing the probability in distance L rather than time. For example, the probability that an electron neutrino is still an electron neutrino after a distance L is given by Equation 2.6.

$$\begin{aligned} P(\nu_e \to \nu_e) = 1 &- 4|U_{e1}|^2|U_{e2}|^2\sin^2\Delta_{21} \\ &- 4|U_{e1}|^2|U_{e3}|^2\sin^2\Delta_{31} \\ &- 4|U_{e2}|^2|U_{e3}|^2\sin^2\Delta_{32} \end{aligned} \tag{2.6}$$

Where Δ_{ij} is $\frac{(m_i^2 - m_j^2)L}{4E} = \frac{\Delta m_{ij}^2 L}{4E}$, E is the energy of the neutrino, and m_i is the mass of the i^{th} mass eigenstate. The oscillation probability depends on the distance traveled by the neutrino, and on the neutrino's energy E. These dependencies mean that a wide variety of experiments, discussed in the following section, can cast light on the mass squared differences Δm_{ij}^2.

2.6.2 Neutrino Mass Measurements

The Equation 2.6 implies that it is possible to experimentally measure the difference in mass between the mass eigenstates and the mixing parameters U_{ij}. There are four main

types of neutrino sources for mass measurement experiments: reactor, beamline, solar, and atmospheric. Solar and reactor sources provide neutrinos at roughly the same energies $\lesssim$ 20 MeV, solar sources have an approximately constant distance L whereas reactor distances can be selected. Atmospheric neutrino distances can be crudely selected using zenith angle, however they have a large range in energies, anywhere from Mega eV (MeV) to Tera eV (TeV). Finally, there are beam experiments, which can provide neutrinos in the Giga eV (GeV) range, and at a selected distance.

One such experiment is the Main Injector Neutrino Oscillation Search (MINOS), which is located in the Soudan mine in northern Minnesota. It works in collaboration with Fermilab, near Chicago, IL, where protons are collided with a graphite target to produce pions and kaons. These particles decay to muons and muon neutrinos (ν_μ) largely, which pass through detectors at Fermilab and the MINOS detector 734 kilometers away through the earth's crust. The ratio in the measured rate of ν_μ between the Fermilab and MINOS detector can be used to measure the mass difference between the second and third neutrino mass eigenstates.

The PMNS matrix is unitary. I.e. the matrix can be fully represented by four parameters corresponding to observable variables. Three variables commonly referred to as mixing angles, and a phase variable. There have been many [36, 39–42] experiments intended to probe mixing angles and neutrino mass differences. The results of these experiments are summarized in Figure 2.5. The figure shows a summary of the values of $\tan^2\theta_{ij}$ and Δm^2_{ij} which have been measured, where θ is one of the three mixing angles. The curves represent the exclusion zones, while the filled in contours represent preferred regions according to different experiments. The borders on the lines specify which two neutrino flavors are being compared. This figure illustrates the necessity for a variety of experiments to determine the neutrino oscillation parameters.

The absolute mass of neutrinos has been constrained by different measurements. The current upper bound measurement for the electron neutrino mass is < 2.05 eV at 95%

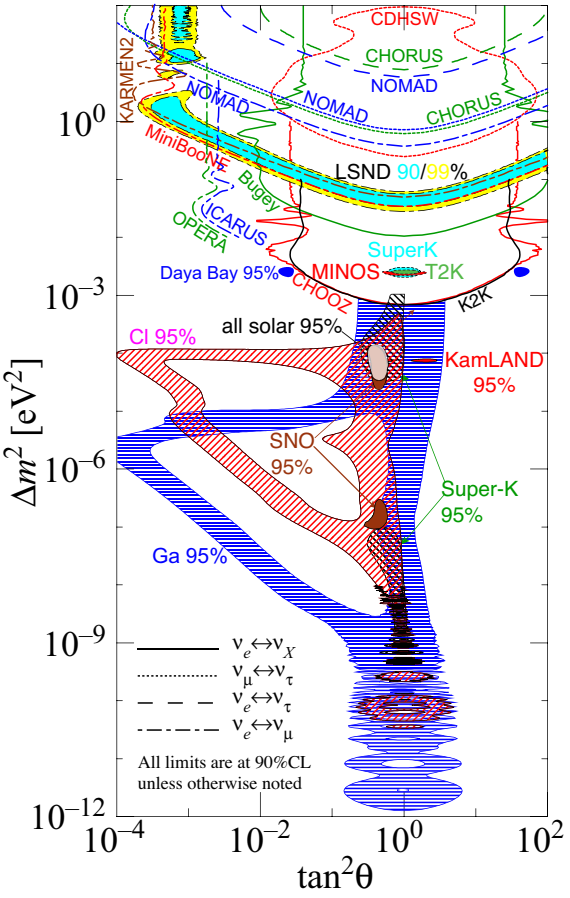


Figure 2.5: Map of exclusion and preferred regions by different experiments of the $\tan^2\theta$ vs. mass difference for neutrinos as of 2014 [43, 44]. Here $\Delta m^2 = m_3^2 - (m_2^2 + m_1^2)/2$, which is positive or negative depending on the mass hierarchy. For this figure the normal hierarchy is assumed.

confidence level performed at the Troitzk experiment [45]. The KArlsruhe TRItium Neutrino Experiment (KATRIN) utilizes the same technique as Troitzk, but hopes to achieve precision below 1 eV [46]. Located in Karlsruhe, Germany, KATRIN takes advantage of Tritium's (H^3) decay to He^3. There are two neutrons and one proton in H^3, one of the neutrons can decay to a proton via a virtual W boson. This particular decay releases a very small amount of energy. At most there is 18.6 keV available for the W boson [47] which instantly decays to an electron and neutrino. There is not sufficient energy available in the decay to create pions or muons or anything heavier. As a result, the electron and

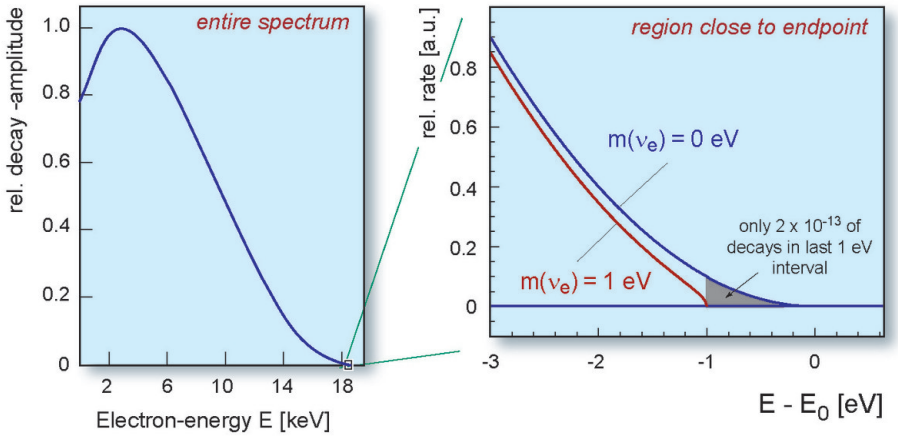


Figure 2.6: $H^3 \rightarrow He^3 + e^- + \bar{\nu_e}$ decay. The figure on the left shows the entire energy spectrum of the electron. The figure on the right shows the electron energy minus the energy available in the decay.

neutrino share the 18.6 keV of energy between them. As the nucleus is much heavier than the neutrino or electron, the majority of the energy is shared by the lighter particles.

Since the neutrino has mass, the neutrino must always carry away at least its mass energy, and the rest of the energy will be carried away by the electron. The interesting case is where the electron carries away the vast majority of the energy in the decay. There will be an upper bound on the electron energy spectrum which is the total energy available minus the mass energy of the neutrino: $E - m_\nu$.

This quantity requires very precise energy measurements and a large number of H^3 decays. The probability of an electron being created with an energy in the last 1 eV of the tritium β-decay energy spectrum is extremely rare, 2×10^{-13} of all decays [48]. The region near the endpoint, shown in Figure 2.6 is the region sensitive to the mass of the neutrino. As a result, this experiment will need to collect data for 5 years in order to have sufficient data [49]. The energy spectrum expected by KATRIN and other Tritium-Helium decay experiments is shown in Figure 2.6.

2.7 Beyond Standard Model

There are several areas where the Standard Model does not accurately or does not at all describe what is seen in nature. These topics are exciting opportunities to discover new physics and fill in the gaps in current understanding of the nature of the universe.

2.7.1 Dark Matter

One of these areas is due to the observation that stars in a galaxy orbit around the galactic center faster than one would expect from the gravitational attraction of visible objects [50]. This could be explained by the presence of massive objects which are not visible, and therefore called Dark Matter. Dark Matter is a type of matter which does not interact via the electromagnetic force, but does through the gravitational force. Neutrinos had been considered an excellent Dark Matter candidate for their "undisputed virtue of being known to exist" [51], however they are not abundant enough to account for the effects seen [52].

There is not currently a good explanation for what this matter might be composed of, but some theories beyond the SM (BSM) suggest a weakly interacting massive particle (WIMP). WIMPs should have a large mass compared to Standard Model particles, which would make them slow moving and classify as cold dark matter. This is a desirable feature, as the evidence suggests that dark matter tends to form clusters rather than exist uniformly in the universe [53].

2.7.2 Matter Anti-Matter Asymmetry

Dirac noticed early in the development of quantum mechanics an equal number of solutions with 'negative energy', which he saw as undesirable [54]. This would imply the Big Bang should have produced both in equal quantities. However, the universe seems to be composed entirely of matter rather than anti-matter.

The only place where anti-matter is detected is in accelerators or due to cosmic rays. The source of this asymmetry is not fully understood. Sakharov proposed three necessary conditions which would produce matter at different rates from anti-matter [55]:

- Baryon number violation
- Charge conjugation violation and Charge-Parity violation
- The universe is in thermal non-equilibrium

CP symmetry is the assertion that the same physical laws hold for a particle when it is replaced with its anti-particle and spatial coordinates are inverted. The P refers to the mirrored symmetry, and the C to the symmetry for particles and their anti-particles. There has been direct evidence of violation of CP symmetry [56], however, the rate at which it has been observed is not sufficient, by several orders of magnitude, to explain matter anti-matter asymmetry of the universe [57].

2.7.3 Neutrino Masses

The Higgs mechanism is the explanation for how the W and Z bosons, quarks and charged leptons acquire mass [58]. However, neutrinos in the Standard Model do not acquire mass from the Higgs mechanism. Equation 2.8 shows the Lagrangian which describes the interaction between the Higgs mechanism and leptons, in this case, specifically the electron. A Lagrangian is an expression of the fields in a system, it is often used in particle physics to describe how different particles interact with each other and to allow for the calculation of branching ratios and cross-sections.

In Equation 2.8 matrix multiplication shows that the neutrino mass and interaction terms disappear, while the charged lepton terms survive. If there was a right-handed neutrino, a minimal extension to the Standard Model could be made, and a term of the form Equation 2.8 could be added to the Higgs interaction, with the e_R replaced with ν_R [58]. This would result in mass terms for neutrinos of the form seen in Equation 2.7.

$$m_\nu = \frac{\lambda_\nu v}{\sqrt{2}} \tag{2.7}$$

However, the Yukawa coupling λ_ν would have to be 10^5 times smaller than λ_e to account for the difference in mass between neutrinos and their corresponding charged leptons [59]. A Yukawa coupling is an interaction between a spin $\frac{1}{2}$ fermion field (Dirac) and a scalar field, such as the Higgs field.

$$\mathcal{L} = -\lambda_e \frac{1}{\sqrt{2}} \left[\begin{pmatrix} \bar{\nu}, & \bar{e} \end{pmatrix} \begin{pmatrix} 0 \\ v+h \end{pmatrix} e_R + \bar{e}_R \begin{pmatrix} 0, & v+h \end{pmatrix} \begin{pmatrix} \nu \\ e \end{pmatrix}_L \right] \tag{2.8}$$

λ_e gives the Higgs coupling to the electron, $(\bar{\nu}, \bar{e})$ is the left-handed lepton doublet, $(0, v+h)$ is the Higgs field, and e_R is the right-handed lepton singlet. Neutrinos sit far below the energy spectrum of all of the other elementary particles. Figure 2.7 shows the masses of the SM quarks and leptons on a logarithmic scale. The wide gap between neutrinos and the other fermions heavily suggests that there might be a non-SM mechanism by which neutrinos acquire mass. The colors in Figure 2.7 refer to the lepton generation, however the neutrinos shown in purple are the mass eigenstates, which mix to form the flavor eigenstates.

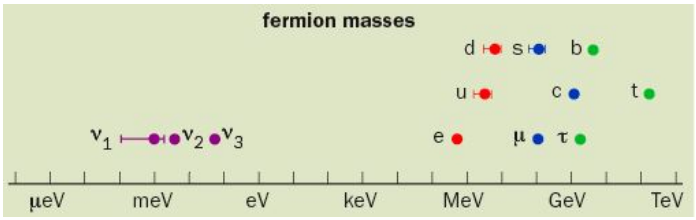


Figure 2.7: Masses of leptons and quarks on a logarithmic scale [60].

An alternative to relying on the Dirac terms to grant mass is to add a Majorana mass term. Majorana particles are particles who are their own anti-particle, whereas Dirac particles and anti-particles are distinctly different. All the fermions in the Standard Model are Dirac.

Majorana couplings essentially allow same-handed particles to couple together. One such mechanism is known as the Seesaw Mechanism. The basis of the name is that the

theory requires a very heavy right-handed neutrino to balance out the very small mass of the left-handed neutrinos already discovered. The Lagrangian for neutrinos, would be of the form in Equation 2.9 [61].

$$\mathcal{L} = -\frac{1}{2}(\bar{\nu_L}\bar{\nu_R^c}) \begin{pmatrix} 0 & m_D \\ m_D & M \end{pmatrix} \begin{pmatrix} \nu_L^c \\ \nu_R \end{pmatrix} + h.c. \tag{2.9}$$

The variables m_D are the Dirac masses which provide left to right handed couplings, while the M is the Majorana mass term between two right handed terms. The masses of the physical neutrino states are the eigenvalues of the mass matrix. This results in a neutrino mass of $m_\nu \approx \frac{m_D^2}{M}$, and a mass of the heavy right-handed neutrino of $m_N \approx M$. The value of m_D is taken to be of similar size to the masses of the charged fermions. The Majorana mass M of the heavy neutrino is then made to be sufficiently large in order to give the neutrino masses observed. Thus M is typically on the order of $\approx 10^{11}$ GeV. This mass is far outside the range of any accelerator. However, the Type III Seesaw mechanism lowers the mass M significantly.

The description above is only the solution for one of the flavor eigenstates. The theory typically requires a heavy right-handed neutrino for each of the light neutrinos.

2.8 Seesaw Mechanism

The Seesaw Mechanism is a minimal expansion of the Standard Model which justifies the smallness of neutrino masses [16]. There are three variations on the Seesaw Mechanism. Type I adds a ν_R singlet per generation, giving neutrino masses of the form [62]:

$$M_\nu = -m_D^T \frac{1}{M_N} m_D \tag{2.10}$$

Here M_ν is the neutrino mass, m_D is the Dirac neutrino mass of the scale of the charged leptons, and M_N is the mass of the heavy ν_R. Equation 2.10 shows the seesaw nature of the mechanism, where the mass of ν_R balances the Dirac masses to give the light SM

neutrino mass. M_ν should be 10^{13} GeV in order to avoid artificially suppressing the Yukawa couplings [63]. Direct tests of the mechanism seem very remote as a consequence of such a large mass.

The Type II Seesaw model instead adds an $SU(2)$ Higgs triplet rather than a single ν_R. Unfortunately, the Type II Seesaw has the same issue as Type I and requires a very high mass scale for the triplets. A natural mass estimate would be $\approx 10^{14}$ GeV [64].

The Type III Seesaw model introduces at least two triplets of heavy particles [65]. However, each neutrino mass can have an associated triplet, implying the addition of up to nine new heavy neutrino particles, three for each flavor. There are variations to the Seesaw model, and even models which contain a mixture of Type I, Type II, or Type III [66–69].

2.8.1 Type III Seesaw Mechanism

The Type III seesaw mechanism [70] introduces at least two extra matter fields in the adjoint representation of SU(2) with zero hypercharge that generate neutrino masses, $M_v^{\mathrm{III}} \sim \Gamma_\nu^2 v^2 / M_\Sigma$. Here M_Σ stands for the mass of the fermionic SU(2)L triplets Σ ,v is the vacuum expectation value of the Higgs field, and Γ_ν is the Dirac Yukawa coupling. The new fermionic triplet components are denoted $(\Sigma^+, \Sigma^-, \Sigma^0)$. Unlike Type I, the Type III Seesaw model couples to gauge bosons that makes studies at the LHC possible. For a single new fermion triplet the Yukawa couplings reduce to a vector: $Y_\Sigma = (Y_{\Sigma_e}, Y_{\Sigma_\mu}, Y_{\Sigma_\tau})$. This analysis utilizes a simplified model and the addition of a single triplet is assumed to only generate a single light neutrino mass, unlike the Standard Model, where there are three. In general one is free to add any number of triplets in order to accommodate neutrino mixing. This model is a low energy limit of a more complete theory. Assuming the addition of a single fermionic triplet has little impact on LHC physics, and therefore is not already constrained by existing measurements.

2.8.2 Phenomenology

Members of the heavy lepton triplet of the Type III Seesaw model, Σ^0 and $\Sigma^{\pm}$, are mainly produced in pairs in the pp collisions at the LHC through gauge coupling, Figure 2.8 [71].

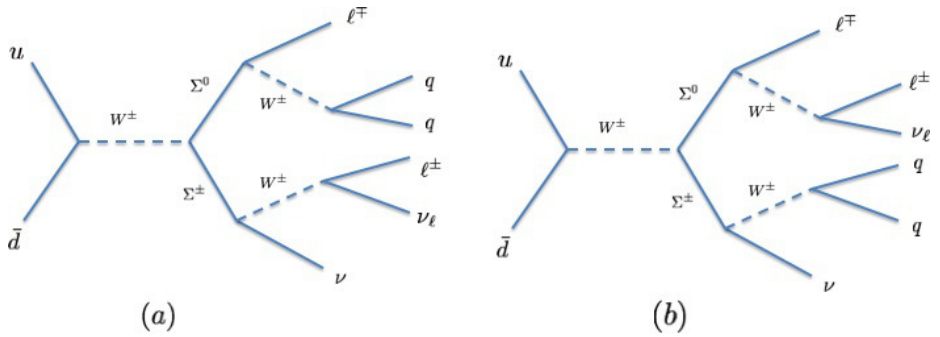


Figure 2.8: Feynman diagram of heavy lepton pair production. Diagram (a) will result in equal likelihood for opposite-sign and same-sign final states. However, diagram (b) will always result in opposite-sign final states. Same-sign means there are two electrons or two positrons in the final state, whereas opposite-sign would have an electron and a positron.

The allowed decay modes of the heavy leptons are listed below:

$$\Sigma^0 \to \nu H$$
$$\Sigma^0 \to \nu Z$$
$$\Sigma^0 \to W^{\pm} l^{\mp}$$
$$\Sigma^{\pm} \to l^{\pm} H$$
$$\Sigma^{\pm} \to l^{\pm} Z$$
$$\Sigma^{\pm} \to W^{\pm} \nu$$

The decay branching fractions depend on the mass of the heavy lepton. They are shown in Figure 2.9.

Taking into account the decay branching fraction, the process with the largest effective cross-section is the one in which both Σ^0 and $\Sigma^{\pm}$ decay to the final states containing a W boson:

$$pp \to \Sigma^0 + \Sigma^{\pm} \to W + \nu + W + l \tag{2.11}$$

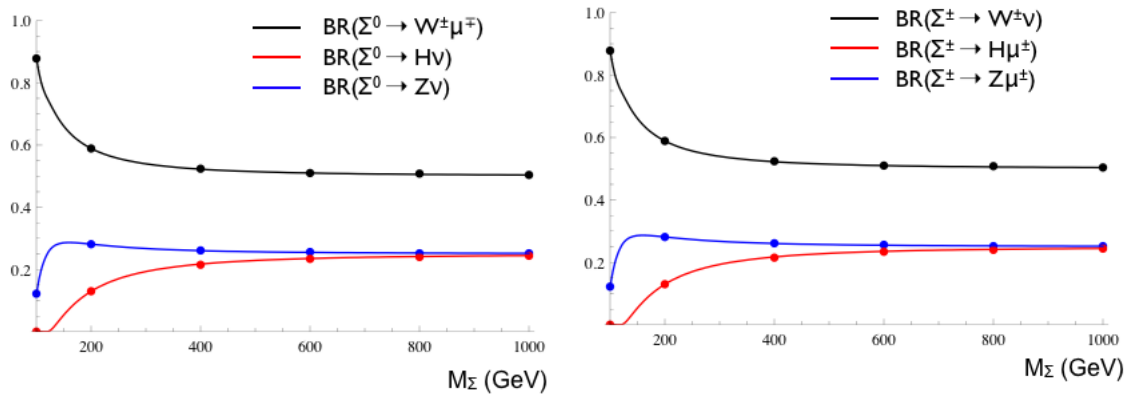


Figure 2.9: Branching ratios of the neutral Σ^0 (left) and charged $\Sigma^{\pm}$ (right) of the heavy fermionic triplet in the case $V_e = V_\tau = 0$, $V_\mu = 0.063$. The dots correspond to numerically evaluated values while the lines correspond to the theoretical predictions [71].

2.8.3 Previous Type III Seesaw Mechanism Searches

There have been several previous searches for the Type III Seesaw heavy lepton. An early search was performed at the L3 detector using the Large Electron Positron collider (LEP), which ruled out charged heavy leptons with masses below 100 GeV [72].

More recently, CMS at the LHC excluded heavy neutrinos consistent with the Type III model from 180 to 210 GeV at 7 TeV [73]. Later in 2015, ATLAS excluded heavy neutrinos decaying into a tri-lepton final state between 100 and 430 GeV [74]. ATLAS also excluded heavy leptons with masses less than 335 GeV, and less than 475 GeV if heavy leptons decay exclusively to W bosons and a lepton [75].

The LHC collider and the ATLAS detectors mentioned here are described in more detail in the following chapter.

2.8.4 Implications of the Seesaw Mechanism

Majorana particles are particles which are their own anti-particles. In a seesaw model with heavy neutrinos which are Majorana particles lepton number can be violated, meaning that neutrino-less double-β decay, Figure 2.10, is permitted. Neutrino-less double-β decay

of neutrons to protons has not yet been observed, despite many experiments. Results from the Germanium Detector Array (GERDA) place a lower limit of 2.1×10^{25} years on the half-life of such a decay in ^{76}Ge [76]. Phase II of the experiment will increase sensitivity to a projected 1.4×10^{26} years [77].

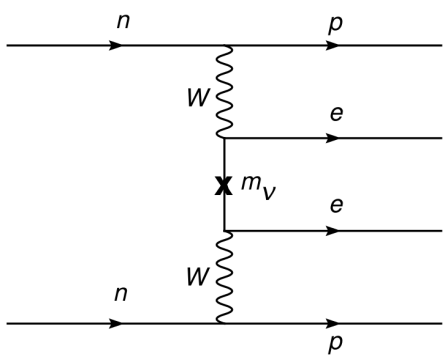


Figure 2.10: Feynman diagram for neutrino-less double β decay.

The Seesaw model can also help solve the baryon asymmetry in the universe. The Seesaw model allows for CP violating phases in the Yukawa couplings between the heavy Seesaw neutrinos and the Standard Model neutrinos. This can result in a preference in decays to leptons rather than anti-leptons [38]. This asymmetry can in turn, contribute to the baryon asymmetry [78].

CHAPTER 3. EXPERIMENTAL SETUP

3.1 CERN

The European Organization for Nuclear Research (CERN) was founded in 1954. It was designed to enable physicists and engineers to probe extreme particle energy ranges at the very limit of what is possible to reach with current technologies. One of CERN's primary goals is to provide particle accelerators for high energy research. These accelerators are used in a wide variety of experiments, with a major focus currently being the LHC. There are also projects targeting application in medical fields, computer science and a wide variety of other topics. CERN has a total of 22 member states. In total, more than 10,000 scientists work and contribute to the organization.

3.2 LHC

The Large Hadron Collider (LHC) is currently the largest accelerator in the world. It is located outside of Geneva, Switzerland and designed for the use of proton-proton and heavy ion collisions at the highest energy regimes ever reached. The LHC is around 27 kilometers in length and sits between 50 and 175 meters below the surface. It is located in the tunnel originally dug for the Large Electron Positron Collider (LEP) which operated from 1989 until 2000, when it was dismantled.

The protons are injected into the LHC at an energy of 450 GeV and are then boosted to approximately 6.5 TeV. The protons gain energy via radio frequency cavities placed around the collider. As protons arrive earlier or later than the timing of the RF cavity, they will be accelerated or decelerated into discrete 'bunches' of particles. There are two beams of proton bunches which move in opposite directions along the beamline allowing for a total center-of-mass energy of 13 TeV at the interaction points. Each of the beams travels in

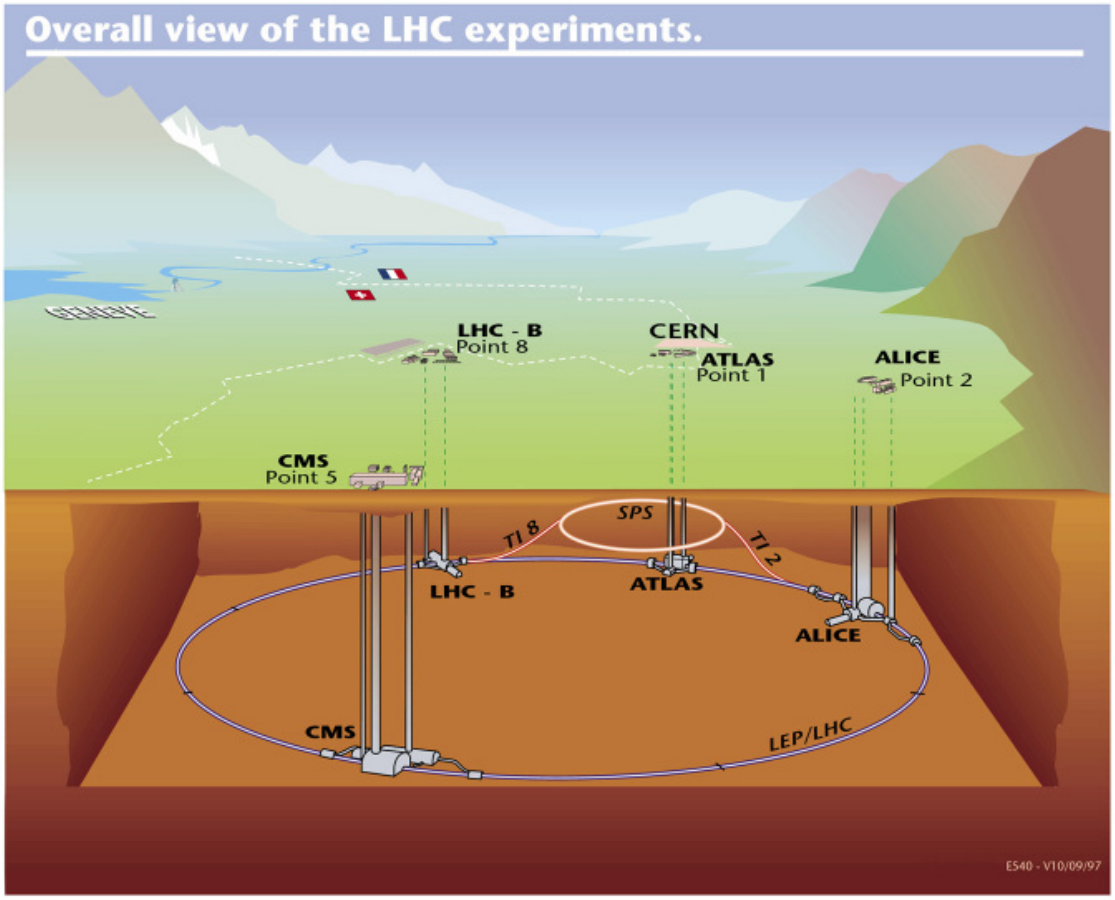


Figure 3.1: Map of the LHC and detector sites.

its own beampipe. The beampipes themselves are held in vacuum, and superconducting magnets are used to turn the proton beams around the loop.

The design collision rate is roughly 40 million bunch crossings per second [79]. There are four main detectors associated with the LHC as seen in Figure 3.1: ATLAS, ALICE, CMS, and LHCb.

The LHC began providing proton-proton (pp) collisions in 2011 at 7 TeV center-of-mass energy. In 2012 the center-of-mass energy was increased to 8 TeV. This was called Run 1 and lasted until 2013. At the date of this publication, the LHC is in Run 2, which began in 2015 and will last until the end of 2018. The analysis described in this document has

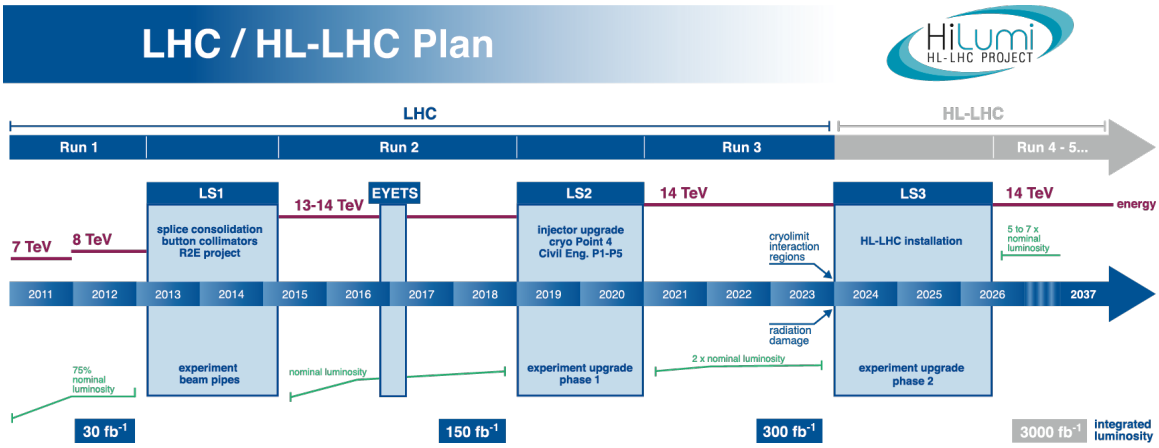


Figure 3.2: LHC long term plans [82].

been performed using the data collected during 2015. The increase in available energy in the center-of-mass frame increases the cross-section for many physics processes, including Higgs production, which goes from having a cross-section of 17.5 pb at 7 TeV to 50.5 pb at 13 TeV [80]. Figure 3.2 shows target energies and integrated luminosities of the LHC project, as well as the scheduled data-taking and upgrade phases. Luminosity is a number of events per unit time divided by the cross section. It can be calculated at the LHC by measuring the number of inelastic interactions per bunch crossing μ, the bunch revolution frequency f_r, and the proton-proton inelastic cross-section σ_{inel} [81]. This quantity is integrated over the number of bunches to give the instantaneous luminosity:

$$\mathcal{L} = n_b \frac{\langle \mu \rangle f_r}{\sigma_{\text{inel}}}$$

Integrated over time, this quantity gives the integrated luminosity of a data collection period.

3.3 ATLAS

The work and analysis described in this document have been performed using the ATLAS (A large Toroidal LHC ApparatuS) detector, which is a large general-purpose detector built

at CERN located at one of the interaction points of the LHC. There are approximately 3000 scientists who are associated with and working with the ATLAS detector, Figure 3.3. It is made up of several sub-detectors, which precisely measure the momentum, energy, and trajectory of various particles. Using these variables, scientists can reconstruct the pp collisions and the subsequent decay of the produced particles to measure the properties and probe the limits of the Standard Model. The ATLAS detector is located below the surface near the border between France and Switzerland. The proposal for the ATLAS detector was submitted in 1994, and construction was finalized in 2008. Each event generates roughly 1.5 MBs of information from the various sub-detectors described below [83]. At a rate of 40 million events a second, this quantity of data is unfeasible to store. Therefore a system of triggers is used to reduce the data to more manageable levels while retaining events which are potentially interesting.

ATLAS is built separated into two main sections: the barrel, and the endcaps. The barrel is in the central region, and the endcaps are the in forward regions. The endcaps are the flat, cap-like ends of the detector. They form a wheel-like shape around the beampipe. This region tends to receive a large amount of particles originating from low-momentum transfer collisions. The barrel is the central cylindrical region. This region is where most of the objects used in this analysis are constrained to originate from. As a result the remainder of the detector description is focused primarily on the barrel.

The barrel generally has fewer scattered particles originating from elastic collisions. As the analysis focuses on the barrel region, only a general description of the endcap detectors will be given.

The ATLAS detector is built centered around an interaction point, where the two proton beams collide. As near as possible to the beampipe, there is a sub-detector called the Inner Detector. Its purpose is to measure the trajectories of charged particles. High precision measurements of the trajectories, especially with measurements close to the beamline allow identification of the original collision location or 'primary vertex'. Directly outside

of the Inner Detector is the Solenoid magnet, which provides the magnetic field that allows momentum measurement within the Inner Detector. The next components moving outwards are the calorimeters. Their purpose is to measure the energies of charged and neutral particles respectively and, in particular, are crucial for jet energy reconstruction. Further outwards is another component of the magnetic system: the barrel toroid magnets. These magnets provide the field necessary for good momentum resolution within the muon spectrometer, which is the outermost detector of ATLAS. In total the ATLAS detector measures approximately 46 meters long with a diameter of 25 meters. The specific sub-detectors within ATLAS are described in detail in the following sections.

3.4 Interaction Point

An interaction point is one of the four points around the LHC where the proton beams are focused and crossed in order to provide collisions. These four points coincide with the locations of the four large detectors. Proton-Proton interactions are caused by passing the bunches of the two proton beams through each other. Each beam consists of bunched of $\approx 10^{11}$ protons. The LHC provides bunch crossings every 50 ns, but part way through the Run 2 operation, this was increased to every 25 ns. On average during the 2015 run there were 15 proton-proton interactions per bunch crossing. If a particular interaction is deemed 'interesting', as discussed further in Section 3.10, the event is recorded. Only a small fraction, $\frac{5}{1000}$[†], of the bunch crossings contain an interesting interaction and will be selected by the trigger. These are interactions with large momentum transfer in the pp collision.

Bunch crossings not selected by the trigger contain largely inelastic collisions with a low momentum transfer. Such collisions produce comparatively low energy physics which have been studied in previous experiments.

[†]With the maximum L1 trigger rate of 100 kHz and spacing of 50 ns, the maximum fraction of 'interesting events is $\frac{5}{1000}$

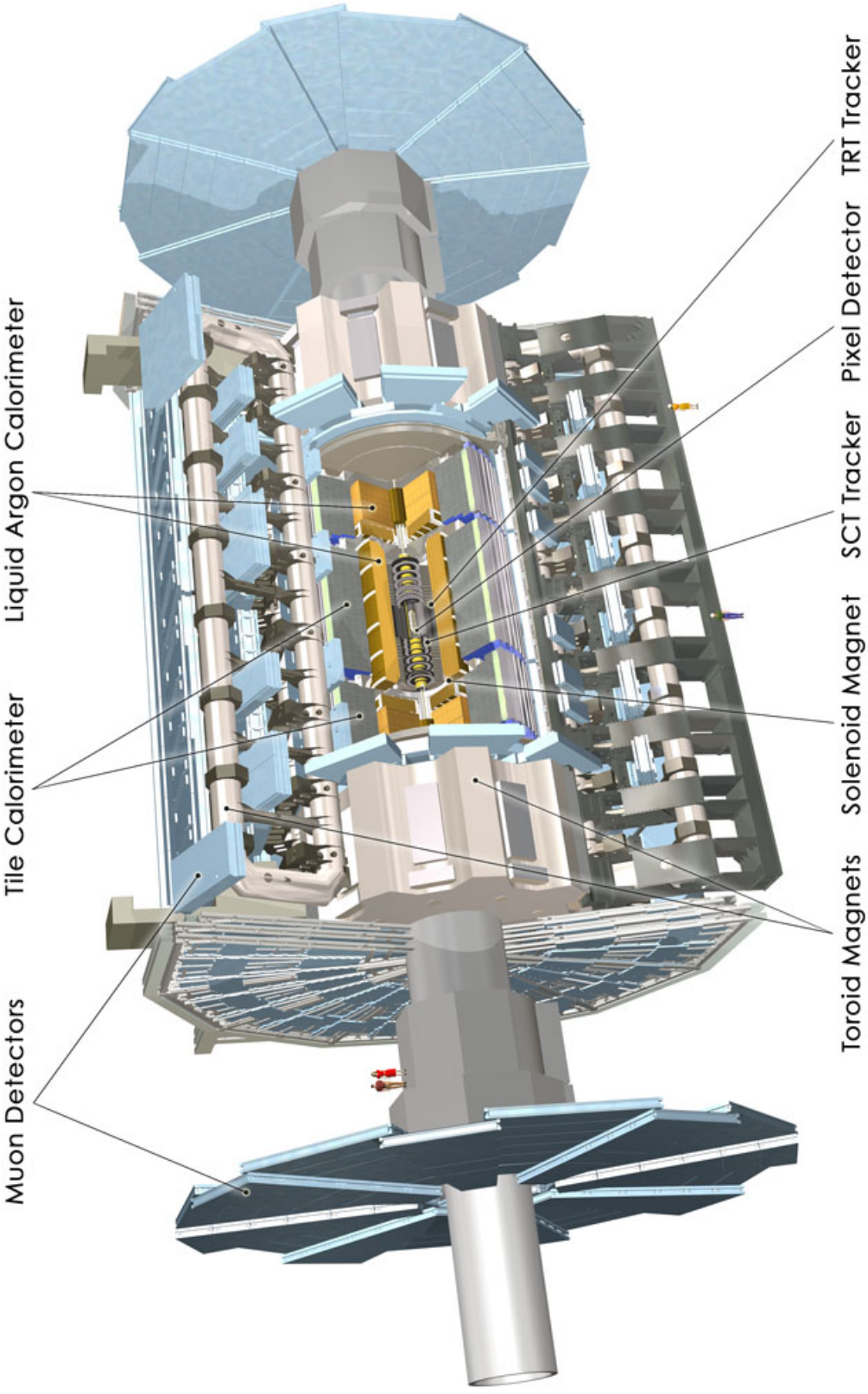


Figure 3.3: A general representation of the ATLAS detector and its subdetectors, with two people on the left side of the main beampipe to give a reference for scale.

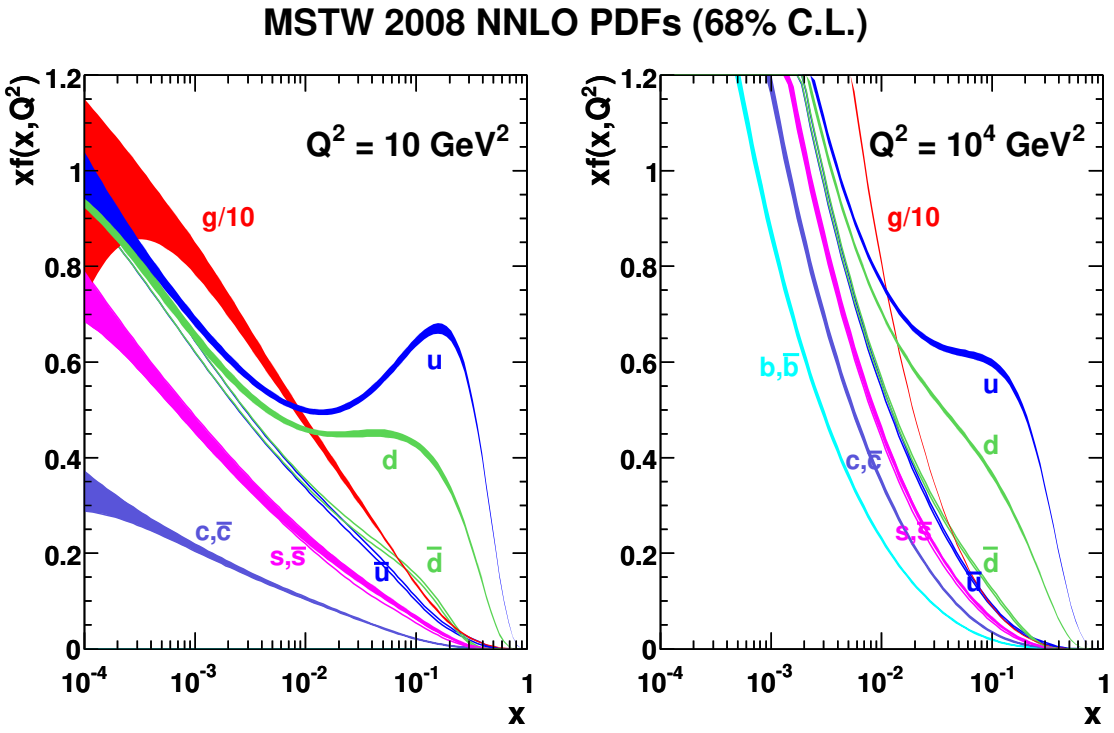


Figure 3.4: Simulation of PDFs, dependent on the portion of the proton's momentum carried by the parton, examined at different energy transfer [85]. The y-axis is a measure of frequency. Gluons are the dominating interacting particle, however, in general at low momentum fraction, sea quarks are the second main contributor. At high momentum fraction the valence quarks are more relevant.

As the protons collide, they do not collide as singular objects. At the energies at the LHC, instead it is necessary to consider the constituent particles, referred to as partons. A proton contains three permanent quarks (two up quarks and one down quark), called valence quarks. The quarks in the proton are held together by gluons, which can spontaneously pair produce quark anti-quark pairs of any flavor. Such quarks are called sea quarks. Depending on Q^2 (momentum transfer squared) the fractions of gluons and quarks with a given flavor vary. In fact, the act of probing the quark content of the proton influences the quark content [84]. Figure 3.4 shows the Parton Distribution Function (PDF) of protons with values of Q^2 relevant for the LHC.

In proton-proton collisions, the full center-of-mass energy is not available to the partons which interact in a collision. This is quite different from electron-positron collisions such as those produced by LEP. New particles can only be produced from the energy available to the partons, which depends on the fraction of momentum which they carry.

The particles produced from proton-proton collisions will either directly pass through the sub-detectors or decay producing secondary particles which will eventually interact with the sub-detectors. The sub-detectors and their components are described in Sections 3.7-3.9. The traces the resultant particles leave allow the reconstruction of the original interaction.

3.5 Coordinate System

ATLAS uses a right-handed coordinate system, where the origin is located at the interaction point in the center of the detector. The x-axis points from the interaction point towards the center of the LHC ring, the y-axis points upwards, and the z-axis points along the beampipe.

For physics analysis, a coordinate system is preferred, where r and ϕ are used in the plane transverse to the beampipe. ϕ is the azimuthal angle around the z-axis. In place of the standard θ, pseudorapidity η is utilized. η is defined in terms of θ as $\eta = -\ln \tan(\theta/2)$, see Figure 3.5. The energy in the transverse plane is defined as $E_{\mathrm{T}} = E\sin(\theta)$. Similarly, the transverse momentum is defined as $p_{\mathrm{T}} = p\sin(\theta)$.

3.6 Magnet System

The ALTAS magnet system consists of three main systems: the barrel toroid, the endcap toroids, and the solenoid. The magnets used at ATLAS are all superconducting, and operate at 4.5 K [86]. These systems work to provide a stable and predictable magnetic field over a large volume. This is crucial for the accurate measurement of the momenta of charged particles as they pass through the tracker systems described in detail in the next few sections. The solenoid sits at a radius of 1.2 meters from the beampipe and is much

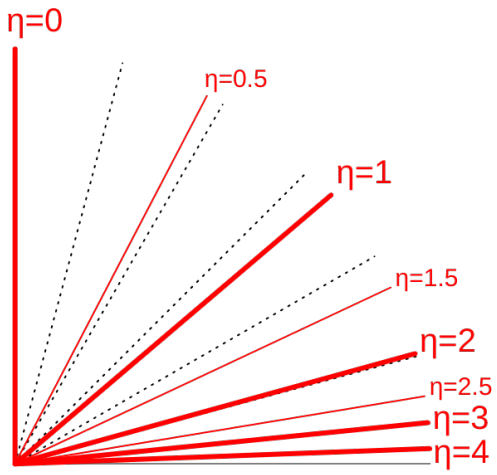


Figure 3.5: Values for pseudorapidity between 0° and 90°. Dotted lines are shown at 15° intervals as a means of comparison.

shorter, at 5 meters in length. The solenoid is designed to provide a two Tesla magnetic field in the tracking region of the Inner Detector to enable momentum measurement and charge identification for charged particles. The barrel toroids are about 26 meters in length and weigh approximately 1300 tons. The barrel toroid is centered around the Interaction Point, and is comprised of 8 separate coils with a peak field of 4 Tesla [86]. These toroids provide the magnetic field which is necessary for the muon spectrometer functionality. Additionally the two endcap toroids have 8 coils each, also with a peak field of 4 Tesla. Figure 3.6 shows the barrel toroids of the ATLAS detector.

Since there are a finite number of coils, the magnetic field is not perfectly uniform [87]. This effect is particularly apparent in the region between the barrel toroid and the endcap toroids, called the transition region. Figure 3.7 shows that the field near the transition region changes significantly.

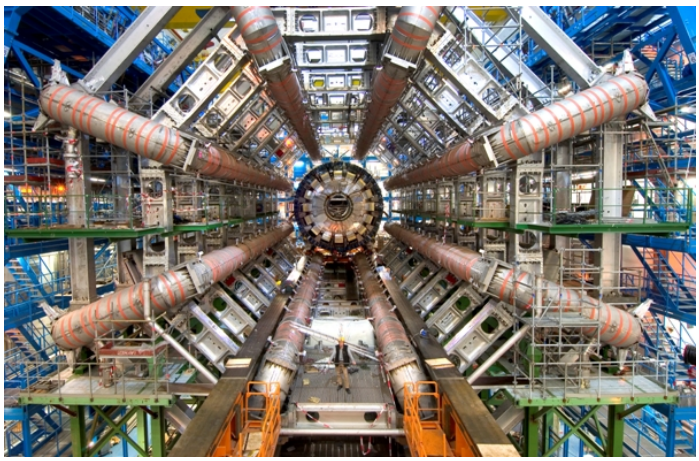


Figure 3.6: The ATLAS barrel toroids. The toroids are the eight objects highlighted with orange. At the bottom in the center is a person for scale.

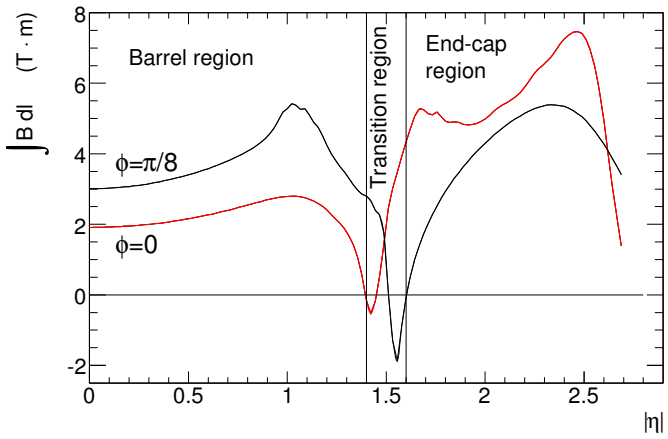


Figure 3.7: Predicted field integral as a function of $|\eta|$ for infinite momentum muons. The integral is calculated for two values of the azimuthal angle ϕ.

3.7 Inner Detector

The Inner Detector (ID) seen in Figure 3.8 is designed to track charged particles through the inner section of the detector. The ID as a whole is approximately 3.5 meters long and has an outer radius of 1.1 meters. It surroundings the beampipe and is centered at the proton-proton interaction point [88]. It sits within a 2 Tesla solenoidal magnetic field described in Section 3.6, such that charged particles will curve as they pass through the field. The curvature creates a helix in the $r - \phi$ plane. As the particles pass through the ID, it creates hits within the three sub-detectors: the Pixel detector, the Semi-Conductor Tracker (SCT), and the Transition Radiation Tracker (TRT).

Using reconstruction algorithms the trajectory of each charged particles in a given event is calculated. The trajectory is used to create an object called a track. The momenta of these tracks can then be calculated based on the curvature of their paths according to Equation 3.1, where p is the momentum, B is the magnetic field, Q is the charge, and r is the radius of the curvature of the track.

$$p = BQr \tag{3.1}$$

3.7.1 Pixel Detector

The Pixel detector is the closest to the beampipe. It consists of 3 layers of pixel modules around the beampipe in the barrel region and 3 disks for each endcap [89]. Figure 3.9 shows the original three barrel region layers of the pixel detector. In 2014, a fourth layer, the Insertable b-layer (IBL) was installed. This layer lies closest to the beampipe and is intended to help better identify b-quarks in interactions. b-quark identification, or b-tagging, is further described in Section 4.5. Each of the pixel layers are 1.4 meters long in the barrel region and centered around the interaction point. The layers consist of 14, 22, 38, and 52 staves, respectively [90]. Each of these staves contains 13 modules, and each module has $\approx 47,000$ pixels. Combined with the endcaps, this totals more than 92 million readout

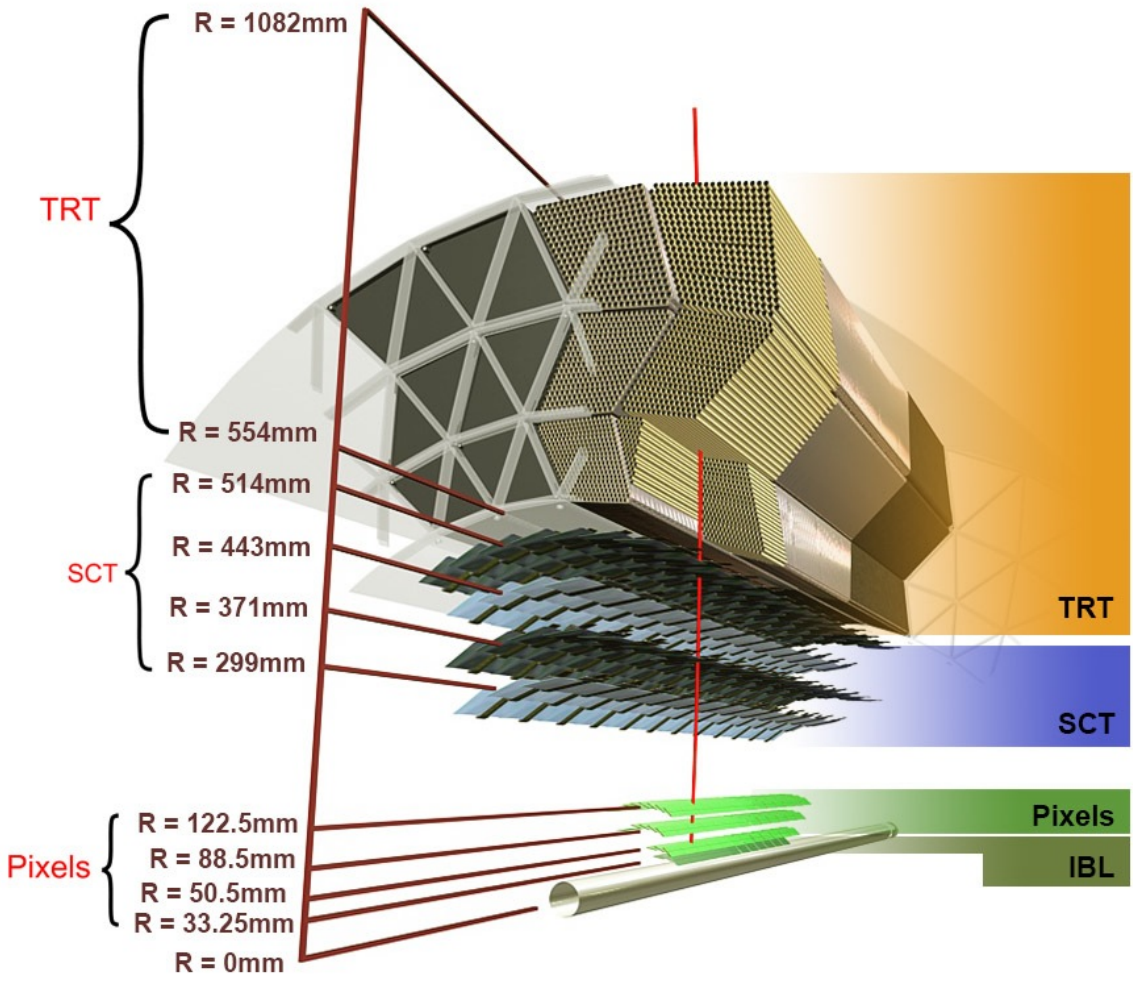


Figure 3.8: Layout of the Inner Detector, comprised of the Pixel detector, the SCT and the TRT.

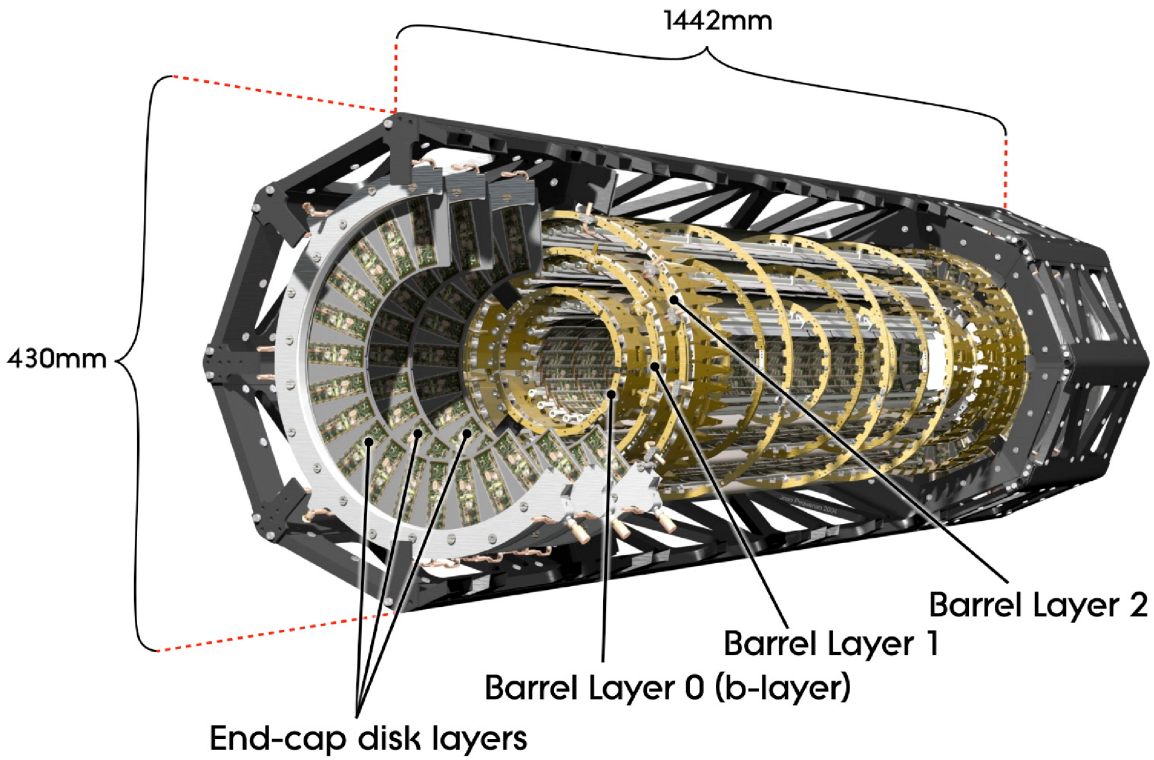


Figure 3.9: Pixel detector.

channels including the addition of the IBL to the Pixel detector. These pixels allow a spacial resolution for hits of 12 µm in $r - \phi$ and 66 µm in z.

The basic physics principle behind the Pixel detector is shown in Figure 3.10. A 'hit' is registered from the following steps: (1) a charged particle passes through the pixel, creating free electrons and holes in the sensor. (2) these electrons and positively charged 'holes' travel to the cathode and anode respectively. (3) The charge flow is registered by reverse bias diodes.

Due to the Pixel detector's close proximity to the interaction point and the high luminosities at the LHC, irradiation is a concern for the electronics involved. Over an extended period of running, radiation damage will cause performance degradation.

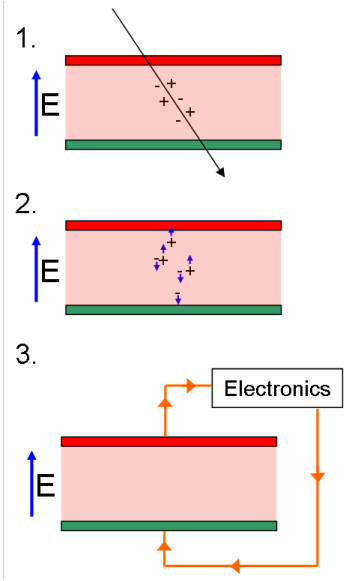


Figure 3.10: Physics of a pixel detector [91].

3.7.2 SCT

The next innermost sub-detector is the Semi-Conductor Tracker seen in Figure 3.11. It consists of four layers of semiconductor modules in the barrel and 9 disks for each endcap resulting in a total of 4,088 modules [92]. The SCT uses the same basic principle for charged particle measurement as the Pixel detector. The arrangement of the SCT allows an average of 4 hits per track. The central barrel gives a coverage of $|\eta| < 1.1$ and the endcaps cover $1.1 < |\eta| < 2.5$. The space-point resolution of this sub-detector is 16 µm in $r - \phi$ and 580 µm in z.

3.7.3 TRT

The outermost sub-detector of the ID is the Transition Radiation Tracker. It is composed of 52,500 axial strawtubes about 1.5 meters in length [92]. The endcaps contain a further 250,000 radial strawtubes. These strawtubes are 4 mm diameter drift tubes, filled with xenon and carbon dioxide gas, each with a central charged wire. As a charged particle enters the strawtube, it ionizes the gas. The ionized particles in the gas travel to the

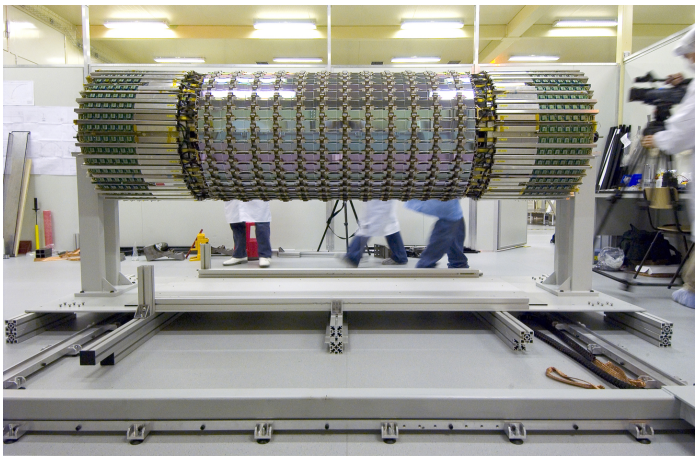


Figure 3.11: The Semi-Conductor Tracker detector before it was installed into the ATLAS detector.

cathode and the freed electrons travel to the central wire which is the anode. The electrical charge is measured and a hit is registered. Due to the orientation of the TRT, this sub-detector only gives $R - \phi$ information. Typically charged tracks will cross more than 36 tubes for particles with $|\eta| < 2$ and $p_{\mathrm{T}} > 0.5$ GeV, and, using timing information, each tube has an intrinsic resolution of 130 μm per straw [92]. In total, there are over 350,000 readout channels in the TRT.

3.8 Calorimeter

The calorimeter, seen in Figure 3.12 is designed to measure the energies of particles. It is divided into two main sections: one for the electromagnetic interactions, and one for hadronic interactions. As highly energetic particles decay, they form 'showers' of particles whose energies can be measured in the detector. An illustration of an electromagnetic shower can be seen in Figure 3.13. A shower is the product of two physical effects: bremsstrahlung (braking radiation) and pair production. As a charged particle interacts electromagnetically with the nucleus of a material it releases energy in the form of a photon. This photon can then pair-produce an electron and a positron. These charged particles in turn create photons, which create more charged particles and so on. This process cascades until it is no

longer energetically possible to do so. Hadronic showers are very similar in structure but involve quarks radiating gluons and pair production of quark anti-quark pairs through the strong force[†].

Bremsstrahlung is proportional to the particle's mass m, the number of protons in the material Z, the energy E of the object:

$$-\frac{dE}{dx}_{\text{Brems}} \propto \left(\frac{1}{m}\right)^2 \frac{Z(Z+1)}{A} Q^2 E \tag{3.2}$$

A is the number of nucleons in the atoms of the absorber material, and Q is the charge of the braking object.

Due to the typically large amount of energy available in the collision, the hadronization from a high energy quark or gluon could spawn hundreds of secondary particles. Rather than treating these particles individually they are treated as a single object, often referred to as a jet of particles. A region, largely conical, is drawn around a cluster of energy deposits in the detector and is treated as a quark or gluon in an analysis. As the hadrons within the jet interact with the hadronic calorimeter they will shower and their energies can be reconstructed. The hadronic calorimeter is a critical tool for analyzing the energies of jets.

3.8.1 Electromagnetic Calorimeter

The Electromagnetic (EM) Calorimeter is designed to identify electrons and photons within an energy range of 5 GeV to 5 TeV. It uses lead as an absorber and liquid argon for the active material. The absorber is a dense material with a large Z, protons in the nucleus, which increases the rate of bremsstrahlung. Figure 3.14 shows the absorber material creating an electromagnetic shower from a single incident particle, which creates many charged particles which enter the active material. The active material measures the resultant

[†]As quarks cannot exist isolated, quarks and anti-quarks bind into mesons. The quarks in the mesons interact with the nucleons of the calorimeter material

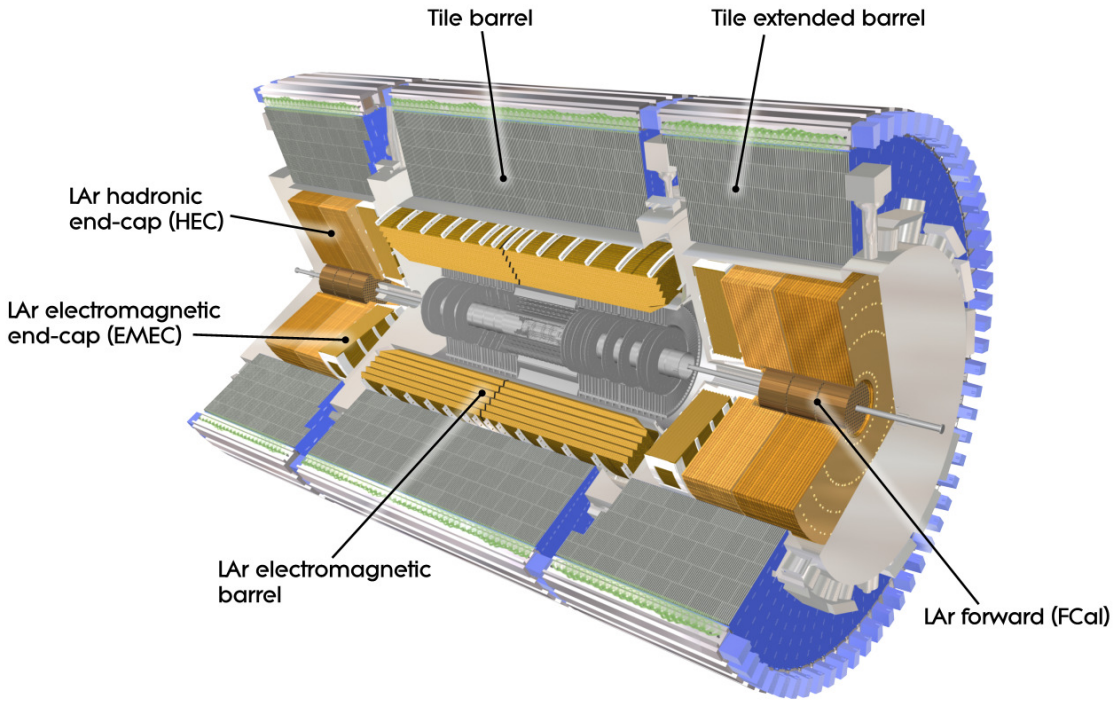


Figure 3.12: Calorimeter schematic.

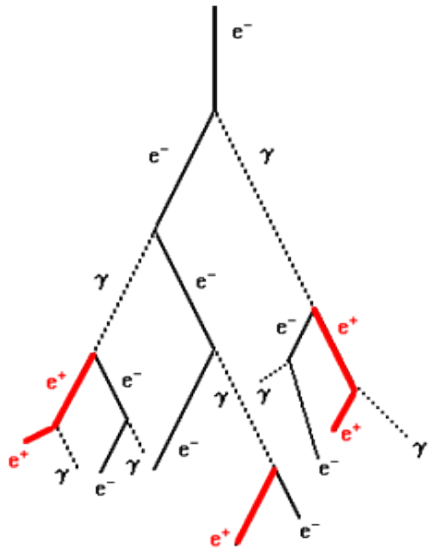


Figure 3.13: An electromagnetic shower from an electron. The red lines show positrons that are produced.

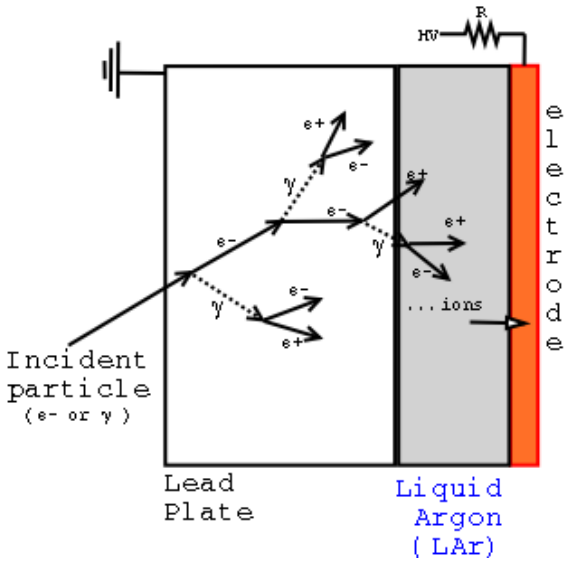


Figure 3.14: An electromagnetic shower within a Lead Liquid Argon calorimeter.

charged particles. The charged particles drift through the liquid argon due to the charge potential in each cell, and the current is registered with the electronics.

The EM Calorimeter is designed with an accordion geometry to avoid azimuthal cracks in the detector, drawn schematically in Figure 3.15 [93]. The accordion geometry reaches up to $|\eta| < 2.5$ and also consists of a pre-sampler up to $|\eta| < 1.8$. The pre-sampler provides information for the triggering of photons and electrons and to ensure a high rate of identification for those objects. Including the endcaps, the EM calorimeter has coverage up to $|\eta| < 4.9$. The EM calorimeter acts as a sampling calorimeter. It is not expected to capture the full energy of all electromagnetic particles which pass through it, but rather measures a fixed fraction of the full energy. The EM shower of electrons and photons are contained entirely within the calorimeter but muons will largely pass through leaving little energy.

3.8.2 Hadronic Calorimeter

The Hadronic Calorimeter is a sampling calorimeter, which is designed to measure a fixed portion of the energy of the hadrons that pass through the detector. The measured values

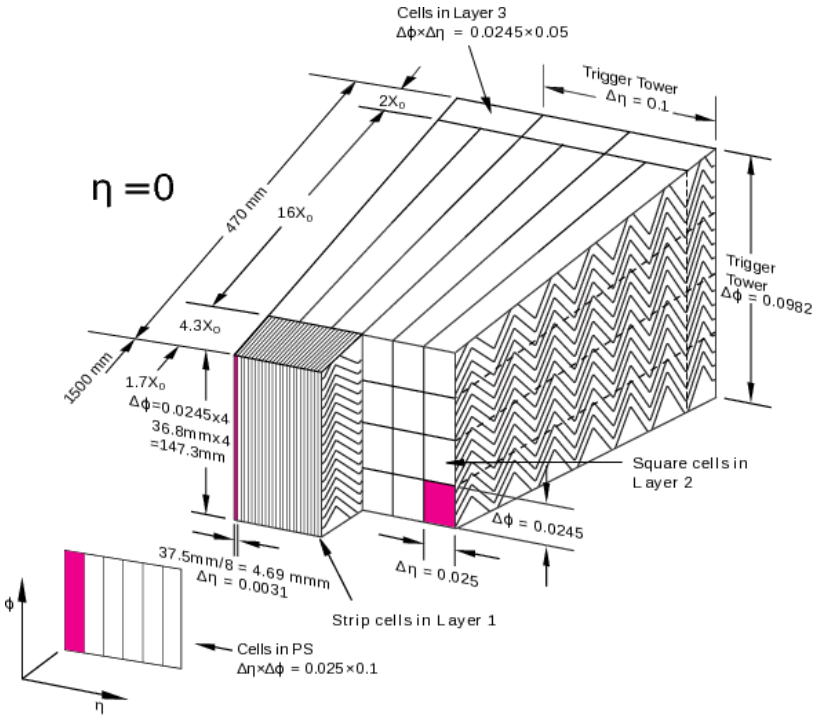


Figure 3.15: Schematic of Electromagnetic calorimeter.

are calibrated and scaled accordingly using known benchmark processes. The Hadronic Calorimeter or TileCal, shown in Figure 3.16, uses iron plates as the absorber and scintillating plastic tiles to produce light from the absorbed energies [94]. The name TileCal refers to these plastic tiles which are used to capture the produced light. High energy hadrons will pass through the EM calorimeter without depositing much energy. The TileCal relies on the strong interaction to produce secondary hadrons from the primary, high-energy hadrons. Charged hadrons will experience bremsstrahlung, and produce photons which will excite the scintillating material. The scintillating material will produce low energy photons which are transmitted via fiber optic cables to photomultiplier tubes to be amplified. The measured photons are proportional to the energy of the incident particle.

The TileCal allows for precise measurement of the energies of hadrons, jets, τ's, and $E_{\mathrm{T}}^{\mathrm{miss}}$. These objects are discussed in more detail in Chapter 4. The fractional energy resolution of the TileCal, σ_E/E, is $\approx 50\%/\sqrt{E}$ [94].

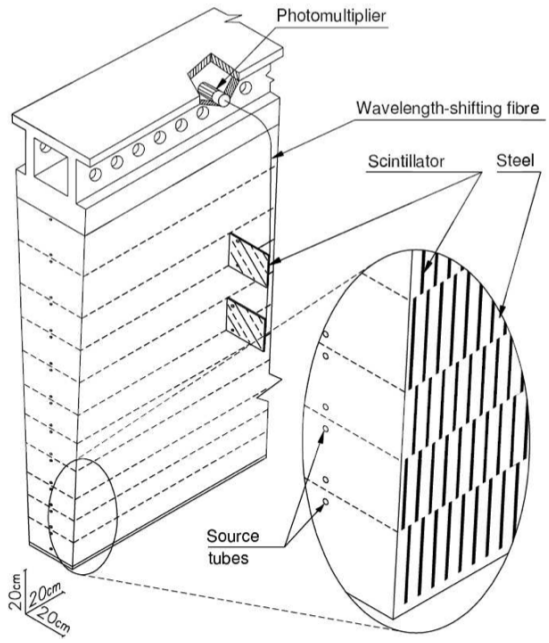


Figure 3.16: Layout of the Hadronic Calorimeter.

3.9 Muon Spectrometer

The muon spectrometer is comprised of four main chamber technologies [87]. The spectrometer is designed such that muons from the interaction point will cross three separate chambers to achieve good momentum resolution. The barrel chambers are located within $|\eta| < 1$ and at radii of $\approx$ 5.0, 7.5, and 10 m. The endcaps cover the region $1 < |\eta| < 2.7$, and are made up of 4 separate disks concentric around the beampipe [87]. There are two styles of detectors in both the endcap and barrel: trigger chambers, and precision chambers. Figure 3.17 shows the layout of the muon spectrometer and the locations of the types of detector used.

The trigger chambers are designed to provide good bunch crossing timing identification for the muons, and to provide a trigger with well-defined momentum cut-offs [87]. This is achieved in the barrel region with the resistive plate chambers and in the endcap with thin gap chambers [87].

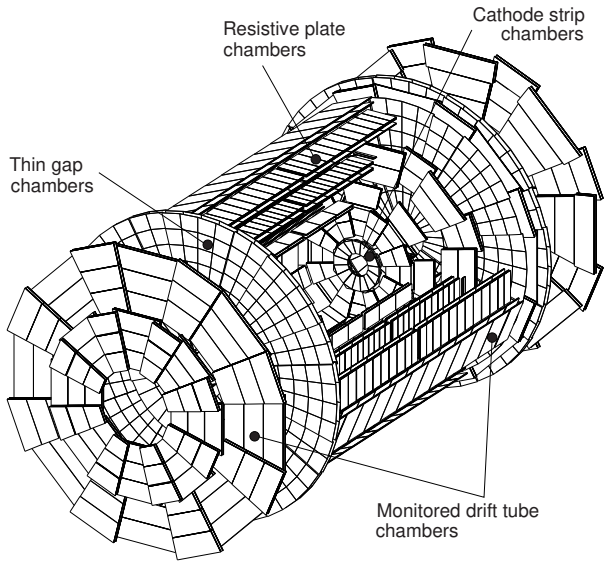


Figure 3.17: Schematic layout of the muon spectrometer.

For the muon precision chambers in the barrel, monitored drift tubes are utilized. Aluminum tubes of 30 mm in diameter are used, filled with an Ar-CH_4-N_2 mixture. The principle for detection is the same as in the TRT. The muons ionize the gas, creating electrons and positrons, which drift through the gas and are registered at the electrodes. The monitored drift tubes are utilized throughout the muon spectrometer except for the higher η region, where the particle flux is very high due to its proximity to the beampipe. Cathode strip chambers are chosen in this region as an alternative [95].

3.10 Trigger and Data Acquisition

The ATLAS detector produces nearly 600 terabytes of raw data every second. It is impossible and unnecessary to store that quantity of data. Additionally, much of the data is from low energy inelastic collisions, which is unlikely to reveal new physics. In order to reduce this staggering quantity of information to more reasonable levels ATLAS employs a

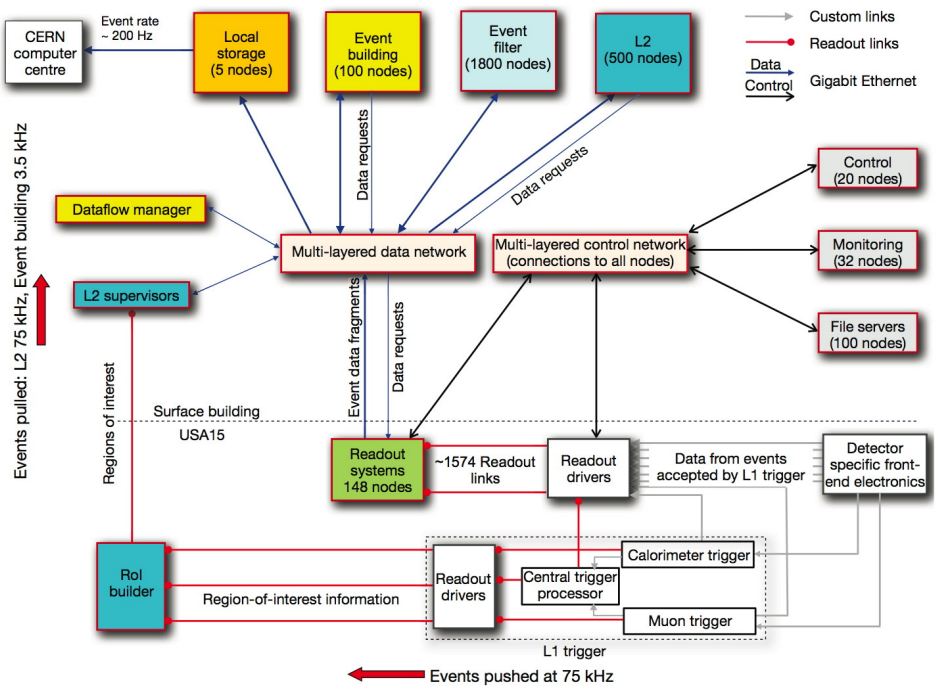


Figure 3.18: Trigger diagram for the ATLAS detector.

system of triggers to select events which may contain new physics. The quantity of data written to disk is reduced to roughly 200 MB/s through the use of the trigger systems. The trigger system at ATLAS was originally separated into three levels, diagrammed in Figure 3.18: Level-1 (L1), Level-2 (L2) and the event filter. The L2 and event filter levels have now been combined to form a single level known as the High-Level Trigger (HLT). The L1 trigger system has been implemented using custom hardware built for very specific purposes. The higher level triggers utilize software largely and are built on commercial CPUs.

3.10.1 L1

The L1 trigger is used to search for indications of high p_{T} simple signatures, such as muons, electrons, photons, and jets in the detector systems. It is also capable of selecting events with large $E_{\mathrm{T}}^{\mathrm{miss}}$(described in Section 4.6), by reconstructing the total available energy in the event. It utilizes primarily the calorimeter information in order to make very quick decisions about whether or not to keep the event. This level of decision making was designed to run at 75 kHz, with the goal of being upgraded to 100 kHz if necessary.

3.10.2 Region of Interest Builder

Events that are selected by the L1 trigger must be passed to the L2 trigger. This is done by collecting Regions of Interest (RoIs) together, and sending them to the L2 where they act as a seed. The RoIs are areas in the detector where the L1 trigger system has identified possible trigger objects, and originate from the calorimeter systems, muons systems and the central trigger processor.

3.10.3 L2 and Event Filter

The L2 trigger uses the RoI information and requests additional information from the relevant readout systems. Utilizing this additional information the L2 trigger makes a decision about whether to keep or reject the event. If the event is kept, all the additional event information is read out, collected, and built into a single event object. This is then forwarded to the event filter, which utilizes the full detector information. Muons and electrons can be reconstructed at this stage. In Run 2, L2 and the Event Filter levels have been combined into a single layer known as the High-Level Trigger or HLT.

3.10.4 Prescaling

Prescaling is a method to reduce the number of frequently triggered events. High p_{T} events tend to have more interesting events with physics that has not been studied at

previous experiments. This data with high p_{T} particles is not prescaled. However, events with lower p_{T} particles occur much more frequently, and analyses in this lower energy range are more likely to be systematically limited than statistically. Writing all available data to tape is impractical and unnecessary, and therefore a prescale is implemented. Only a fraction of the low p_{T} events are recorded to avoid overwhelming the bandwidth of the trigger system and disk-write speed. Such events are scaled by a weight to adjust to the quantity that was actually produced. Scaling is also used to do similar adjustments to Monte Carlo simulations to create good agreement with recorded data. These simulations are discussed further in Section 5.2.

CHAPTER 4. OBJECT RECONSTRUCTION

The ATLAS detector has been designed to detect and measure Standard Model particles resulting from the proton-proton collisions delivered by the LHC. This chapter focuses on the particles produced in collisions at the LHC and how they manifest in the ATLAS detector. The Type III Seesaw Mechanism model investigated has a specific final state which is described in more detail in Chapter 5. Final state refers to the signature of a particular process. In the context of this analysis at ATLAS, the relevant particles include electrons and muons. Additionally hadronic jets and the negative sum of the momenta in the event are necessary to fully reconstruct the events consistent with the signal model.

4.1 Pile-up

The cross-section for inelastic pp interactions at the LHC is expected to be around 67 microbarns. With this cross-section, and 10^{11} protons per bunch, and with the luminosity in the 2015 run at the LHC, a bunch crossing has an average of 15 interactions. This can be problematic for analyses, since any recorded event in the detector contains a collection of particles coming from several different collisions. The term for these additional interactions is pile-up, specifically in-time pile-up. In addition, out-of-time pile-up occurs when events from previous bunch crossings bleed into the current crossing. This can happen since the time between crossings is only 25 ns, and the response time of some detector components can be much longer. These pile-up interactions are typically low p_{T} interactions producing low energy jets. A challenge for analyses which use information from the calorimeter is reducing the effect of energy deposits from the additional, low p_{T} interactions in a given bunch crossing.

4.2 Primary Vertex

There are many tracks associated with any given bunch crossing of protons. These tracks are traced back to the center of the detector and used to identify primary vertex candidates. A primary vertex is the vertex which contains potentially interesting physics. The primary vertex candidates are ordered according to the sum of p_{T}^2 of tracks in the vertex. It is common to take the vertex with the largest sum of p_{T}^2 to be the primary vertex.

4.3 Leptons

The three charged leptons of the SM and their associated anti-particles interact differently with the material of the detector due primarily to their masses. In general, electrons are expected to be stopped in the EM calorimeter, and their energy well measured. As they travel through the Inner Detector, they will ionize material and leave a track of their path.

Muons will pass through most of the detector and will not be stopped by the calorimeters, but will still ionize material. With the Muon Spectrometer muons can easily be identified correctly and their momenta correctly measured.

Taus will decay virtually at the vertex, their lifetime is sufficiently shorter, that they will not reach the innermost subdetectors. They might end up being reconstructed in their decay to one or three charged mesons or to one of the lighter leptons, depending on whether they decay leptonically or hadronically. The most common decay modes of the τ lepton are shown in Figure 4.1 [96]. The hadronic mode is available to the τ, as the quarks from the W boson will have enough energy to form pions or kaons. Whereas with the muon, the hadronic decay mode is not available, as the muon mass is 105 MeV and the lowest meson mass, the pion, is 140 MeV.

Taus decay hadronically 65% on average, meaning they will be reconstructed as a low track multiplicity jet. The remainder of the time they will show up as a lepton in the detector, with equal likelihood to decay to an electron or muon. However, some portion

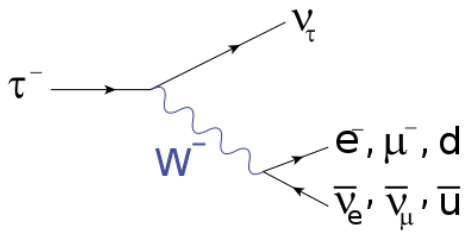


Figure 4.1: Tau lepton decaying to W boson and τ neutrino, the W is able to decay hadronically here [96].

of the total energy of the τ will have been lost to the neutrino(s) produced, making it impossible to fully reconstruct the original energy.

4.3.1 Electrons

Electrons are identified by using the Inner Detector and Calorimeter energy depositions to reconstruct the electron's energy. Electrons are produced at the LHC within a momentum range between 4.5 GeV and 1 TeV [97]. These electrons are fully captured within the electromagnetic calorimeter, meaning their shower products will not pass through the calorimeter. This is important since the total energy of the electron would not be fully represented by the deposits in the calorimeter, causing problems with reliably reconstructing events. Events with electrons are preselected using triggers as discussed in Section 3.10. For this analysis, the lowest unprescaled triggers were chosen to initially select events. For the 2015 dataset collected by the ATLAS detector, these are the triggers that were appropriate:

- HLT_e24_lhmedium_L1EM20VH
- HLT_e60_lhmedium
- HLT_e120_lhloose

The events in these triggers were required to have at least one high energy electron, so-called single electron triggers. The first term refers to level of the trigger, Level 1 ,or the high level trigger. The second term clarifies what object and p_{T} was needed to satisfy

the trigger. e stands for electron, and 24, 60, 120 is the p_{T} in GeV. The final term is the stringency of likelihood algorithm used to identify the electron. Additionally, the lowest p_{T} trigger specifies the L1 trigger which was passed, which originated from the EM calorimeter with a minimum p_{T} of 20 GeV.

Two collections of electrons are defined. One being *loose*, for the purpose of 'fake' electron estimation, and a subset of the loose, called *tight*. Loose and tight refer to how stringent the selection criteria are on electrons. The tight definition is necessary for electrons used in the final state to reduce contamination from jets being reconstructed as electrons or charge mis-identification. However, it should be noted that the loose selection is already quite stringent, rejecting over 99% of background candidates [98].

Several cuts, listed below, are used to specify electrons with from the possible electron candidates. This is important to get a collection of electrons with high purity. All electrons in this analysis are required to pass these cuts. One of these is a simple p_{T} cut, which ensures that the electrons pass the low p_{T} trigger requirement with a well-understood efficiency, Figure 4.2. The trigger efficiency decreases sharply from 30 to 20 GeV.

Another requirement for electrons is that they lie within well measured parts of the detector. There is a region called 'the crack', which is the border between the barrel and endcap regions of the detector. Electrons in this region are excluded from the analysis, as they will not have well reconstructed energies. Additionally, there are a large number of particles at large values of η, from low p_{T}-transfer scattering, making it difficult to reconstruct objects accurately. The additional partons from protons in a collision are more likely to produce forward jets, making measurement in the endcap regions difficult. Electrons in this region are also excluded.

Furthermore, there are a few cuts on variables derived from track parameters ($\frac{d_0}{\sigma(d_0)}$, $|z_0 \sin(\theta)|$) to ensure that the electron matches to the primary vertex. There is also a cut called LooseLH, which utilizes a number of parameters, including the number of hits in different Inner Detector sub-detectors to create a likelihood that a given object is an

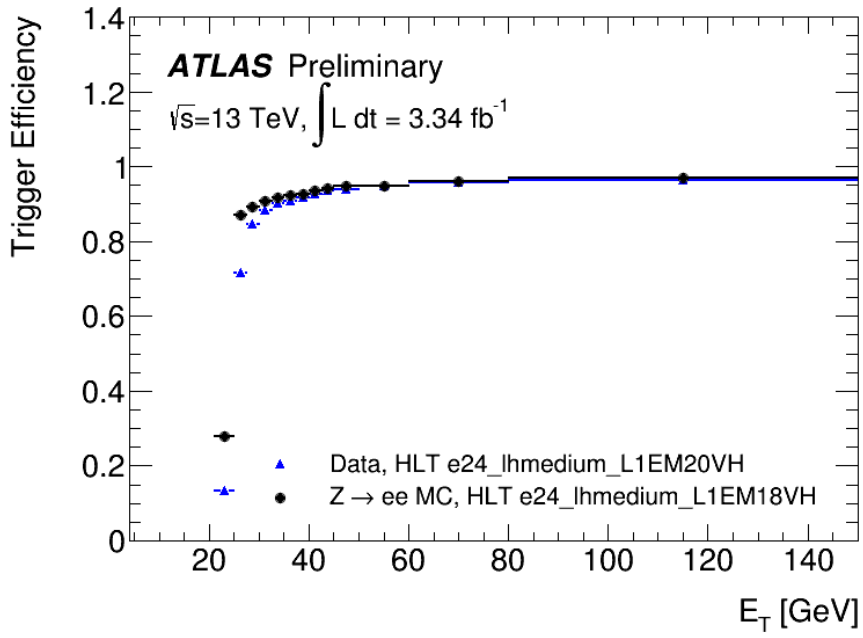


Figure 4.2: Trigger Efficiency Curve, or trigger efficiency turn-on curve for single electron HLT in data and MC simulated events.

electron, and that it originates from the IP. This likelihood is a discriminant which is separated into a few different levels of identification.

This summarizes the cut requirements for loose electrons:

- $p_{\mathrm{T}} > 30$ GeV/c
- $|\eta| < 2.47$ and outside the crack: $1.37 < |\eta| < 1.52$
- $\frac{d_0}{\sigma(d_0)} < 5$
- $|z_0 \sin(\theta)| < 0.5$ mm
- LooseLH [98]

An isolation requirement and tighter identification was also required to define the tight electrons. Isolation is a method to help distinguish electrons from jets, as they can appear to be very similar in the calorimeter. The level of isolation is calculated using cones around the

electron to measure the total p_{T} and E_{T} deposition around the object within a certain radius from the electron's trajectory. Cuts are defined on these isolation parameters dependent on the η and p_{T} of the electron. Working points are defined at certain target efficiencies. For the case of Loose isolation, which was used in the analysis, the efficiency is $\approx 98\%$.

Summarizing, tight electrons further require:

- MediumLH [98]
- Loose isolation [99]

4.3.2 Muons

Muons are easier than electrons to identify and to reconstruct in the ATLAS detector. They pass through the Inner Detector, causing hits which can be used to reconstruct a track. They then pass through the Electro-magnetic calorimeter, depositing very little energy. Equation 3.8 shows that a muon will lose 44,000 times less energy due to bremsstrahlung when compared to an electron of equivalent energy. The muon will pass through the remaining detector and cause hits in the muon spectrometer. From the combined track data, the muon momentum can be measured with high precision; $\approx 2.5\%$ for muons with $p_{\mathrm{T}} = 25 - 75$, growing to 11% at $p_{\mathrm{T}} = 1$ TeV [87].

Muons are not directly used in this analysis, instead they are used to veto events. They pass some ID requirements specifying a certain number of hits in the pixel, SCT and TRT detectors. Loose quality muons are required in the analysis. Depending on which subdetectors identified a muon, different cuts are performed [100]. The strongest discriminant between muons and other physics objects is how far they pass through the detector. The muon spectrometer is critical to identify and correctly measure the momentum of muons at ATLAS.

Additionally the analysis requires that any muon has a minimum of 7 GeV/c for p_{T}. The final requirement is that the muons are very loosely isolated. The method for calculating isolation is the same as described in the electron section. The main difference for the muons

is that the only the tracks are utilized to calculate isolation for LooseTrackOnlyIso. The combined requirements to define muons used in this analysis are summarized below:

- Loose quality [100]
- $\frac{d_0}{\sigma(d_0)} < 5$
- $|z_0 \sin(\theta)| < 0.5$ mm
- $p_{\mathrm{T}} > 7$ GeV/c
- pass LooseTrackOnlyIso [99]

4.4 Jets

Jet are reconstructed using various algorithms to collect energy deposits in the calorimeter together in a way such that the objects contained within a given jet are likely to have come from the same jet. The standard algorithm used for analysis is anti-k_t [101], and the typical size of jets used at ATLAS is a radius of 0.4 in $\eta - \phi$ space, see Figure 4.3. The radius R is defined in Equation 4.4.

$$R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{4.1}$$

Jet reconstruction at the LHC is difficult because it is very sensitive to pile-up. There are various cleaning methods that are used to reduce the pile-up contribution to jet energies.

Figure 4.4 shows a di-jet event within the ATLAS detector. The multi-colored curved and straight lines represent tracks left within the Inner Detector. The event shows two main jets, one with energy ≈ 85 GeV, and the other of about 30 GeV, as seen on the right. The circles represent the distinct jets that the deposits in the calorimeter are grouped into. Although there are six total jets, the remaining four are likely pileup effects, as they are much lower energy jets. This particular event is from the 2010 data collection, where there

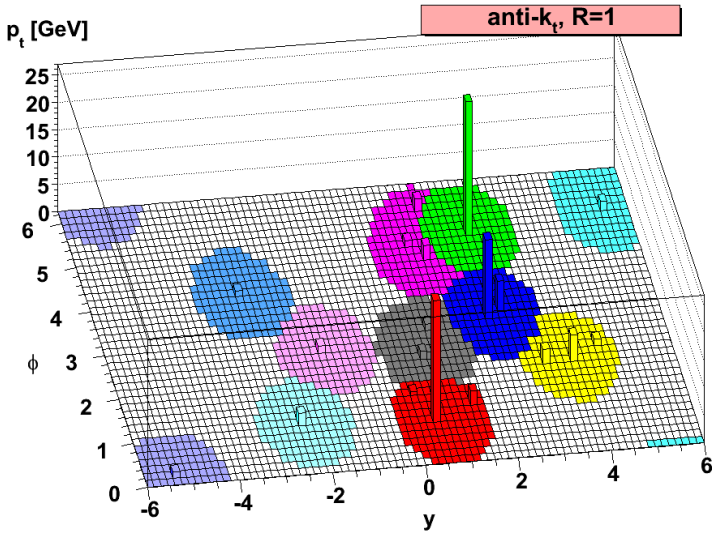


Figure 4.3: A sample parton-level event with jets reconstructed with a radius of R=1.0. The reconstruction uses the anti-k_t algorithm. [101]

were approximately five interactions per crossing [102]. In the other two sections of the figure, the tracks can be clearly seen, showing the curvature of charged particles passing though the magnetic field in the detector and interacting with the sub-detectors.

To ensure that the jets that are used in the analysis are well reconstructed, several cuts are applied. A momentum cut of 20 GeV is used and a requirement is applied that the jet is in the central region of the detector ($|\eta| < 2.5$), rather than the forward segments. The standard method used at ATLAS for eliminating jets from pile-up events is the Jet Vertex Tagger (JVT) [103, 104]. The JVT is a discriminant constructed using Jet Vertex Fraction (JVF). JVF is a variable used to identify the primary vertex of a specific jet and takes advantage of tracks from the Inner Detector to do vertex matching [105]. JVF is the ratio of the sum of p_{T} of matched tracks which originate from the chosen primary vertex to the sum of p_{T} of all matched tracks in the jet, independent of their origin, see Equation 4.2.

JVT uses the JVF and $R_{p_{\mathrm{T}}}$ to calculate a likelihood that a given jet is a jet from the primary vertex [103]. $R_{p_{\mathrm{T}}}$ takes a scalar sum of the p_{T} of tracks associated with a jet divided by the calibrated and pile-up subtracted jet p_{T}.

$$\mathrm{JVF}(\mathrm{jet}_i,\, \mathrm{PV}_j) = \frac{\sum_k p_{\mathrm{T}}(\mathrm{track}_k^{\mathrm{jet}_i}, \mathrm{PV}_j)}{\sum_n \sum_l p_{\mathrm{T}}(\mathrm{track}_l^{\mathrm{jet}_i}, \mathrm{PV}_n)} \tag{4.2}$$

The pile-up energy contribution is subtracted using an area-based method [105]. The pile-up energy contribution is calculated from the average pile-up energy density in the $\eta \times \phi$ plane and the area of the jet in the plane [106].

Besides pile-up, there are several other sources of non-primary vertex jets [107]. One source is through cosmic-ray interactions. This can cause muons to reach the detector and leave depositions within the calorimeter cells at energies which can be similar to typical jets from the pp collision. Another source can be from proton collisions upstream of the detector. Again these collisions can produce muons, again producing calorimeters depositions. Finally, some events need to be rejected due to large electronic noise in the calorimeter readout. Persistently noisy cells are masked in event reconstruction, but sporadically noisy cells are masked on an event-by-event basis. The following summarizes the requirements for jets in this analysis:

- Jet $p_{\mathrm{T}} > 20$ GeV
- Jet $|\eta| < 2.5$
- Pass JvtMedium cut [108]
- Pass Jet cleaning cut [109]

4.5 b-jet Tagging

b-tagging refers to the methods used to identify jets which are created from the decay of b quarks. The existence or absence of b quarks in a particular event can be very useful for rejecting backgrounds. For example, a top quark will nearly always produce a b quark as it decays, as shown in the CKM matrix in Section 2.4.3, however the Type III Seesaw signal model will not produce any b quarks.

There are several means of identifying b quarks. One of the main mechanisms is due to the B hadron's relatively long lifetime. For example, $B_d^0, B_s^0, B^{\pm}$ mesons decay an average of 3 mm away from the primary vertex, creating a secondary vertex that is displaced from the primary vertex [61]. Therefore, tracks from the Inner Detector can be traced back to the two vertices and can be used to identify a B hadron within the jet of particles. Figure 4.5 shows how the variable d_0 can be used to associate tracks from different vertices in the same event. The d_0 variable is one of several variables used to identify b-jets in the ATLAS detector. The current method utilizes multivariate techniques and a large number of variables to improve the ability to identify the b-jets [111].

4.6 $E_{\mathbf{T}}^{\mathbf{miss}}$

When neutrinos are produced in an event they pass through the detector without leaving any tracks or any energy deposition in the calorimeters. Neutrinos can be detected by examining the conservation of momentum in a particular event. The total momentum of the initial pp system in the $x - y$ plane is essentially zero with respect to the magnitude of the momenta involved in the collision. When a neutrino is produced in a collision, the p_{T} of the produced particles will not sum to zero, as the neutrino will not be able to be detected. The ϕ direction of the neutrino's path can also be determined from the negative vector sum of the total p_{T} in the collision. The magnitude of the negative vector sum of p_{T} is referred to as the missing energy, or $E_{\mathrm{T}}^{\mathrm{miss}}$, since the neutrino's mass is so small.

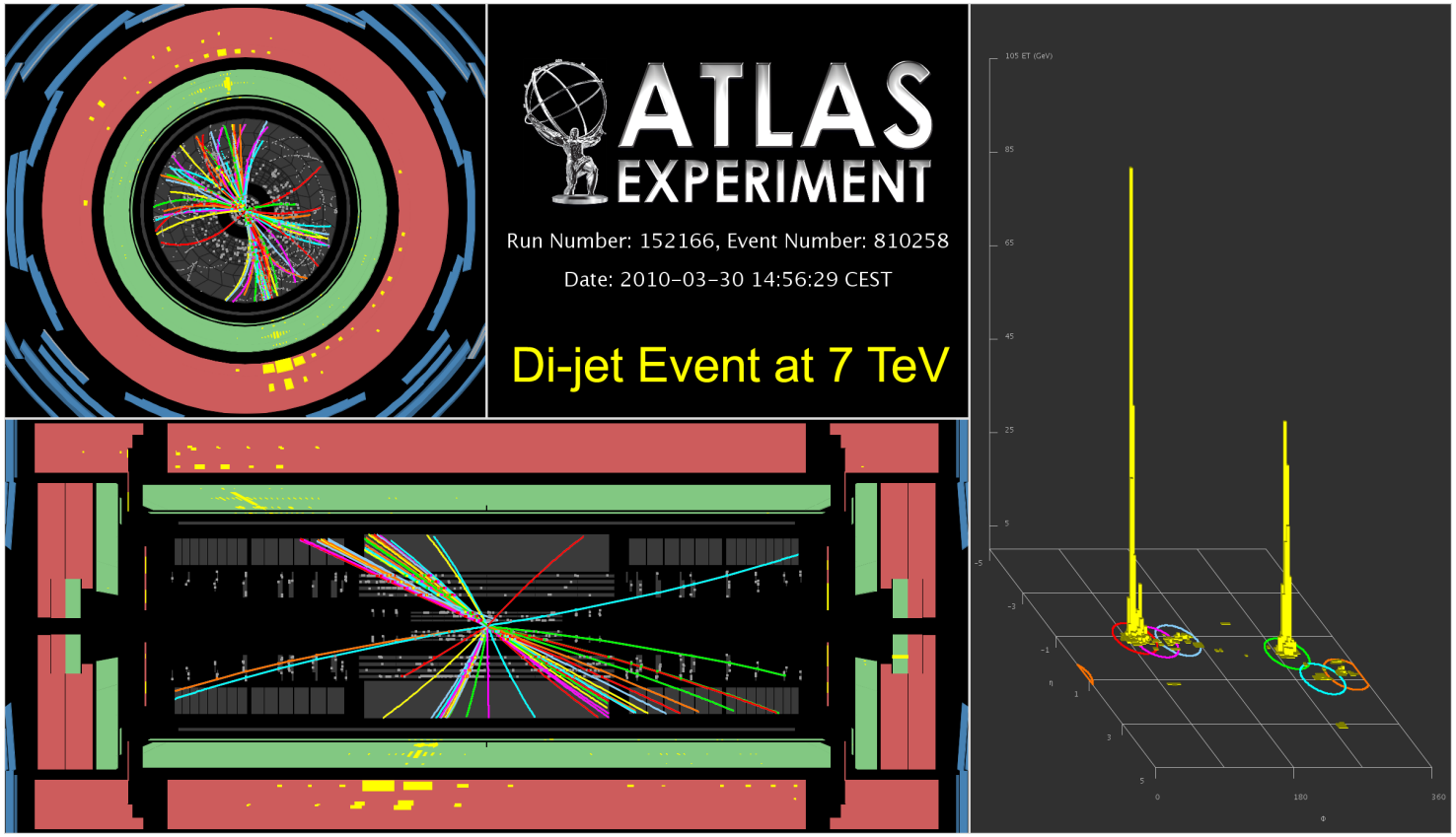


Figure 4.4: A di-jet event at the ATLAS detector. The Inner Detector is depicted by the gray regions. The EM calorimeter is represented by the green regions, while the red regions are the Hadronic calorimeter. Finally the region in blue is the muon spectrometer. In the top left, the $r - \phi$ plane is shown. Bottom left shows the $r - z$ plane, and the right plot shows the energy deposition in the $\phi - \eta$ plane. Six jets are reconstructed, two of which have large energy deposition in the calorimeter.

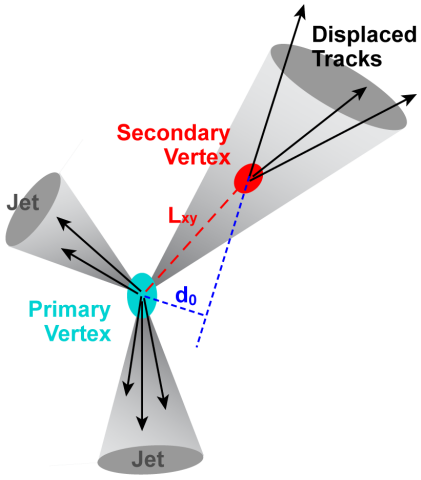


Figure 4.5: Displaced secondary vertex from a decaying B hadron [110].

CHAPTER 5. SIGNAL AND BACKGROUND DESCRIPTION

This chapter details the signal and background contributions in the signal region where the signal cross-section measurement is made. Region is a term defining a parameter space. For example, the Z boson region would be the parameter space where Z bosons are selected and the contributions of other physical processes are reduced. Typically, analyses use Monte Carlo generated data in order to simulate contributions from known physical processes. The Monte Carlo is used to measure the signal efficiency in order to make a measurement for the signal cross-section and set a limit or record a discovery. In this analysis, the dominant backgrounds cannot be estimated using Monte Carlo and must be derived with data-driven methods. These dominant backgrounds originate from two effects: electron charge mis-identification, and jets which fake electrons. The sub-dominant background processes are estimated using Monte Carlo.

5.1 Final State and Analysis Strategy

Final state refers to the particles measured by the detector for a certain physical process. The cross-section measurement is performed in the final state that both $\Sigma^{\pm}$ and Σ^{0} decay to states containing a W boson. One W boson decays leptonically and the other W boson decays hadronically. The Feynman diagram of the production and decay process searched for in the analysis is shown in Figure 5.1. As a result, the signature of the final state in the analysis contains two high p_{T} leptons, two jets from the hadronically decaying W boson, and large $E_{\mathrm{T}}^{\mathrm{miss}}$. Note that in this final state the two high p_{T} leptons can have the same-sign (SS) or opposite-sign (OS) electric charged and, at the same time, they can also be of same flavor (ee or $\mu\mu$) or different flavor ($e\mu$). τ's are also considered in the analysis, however,

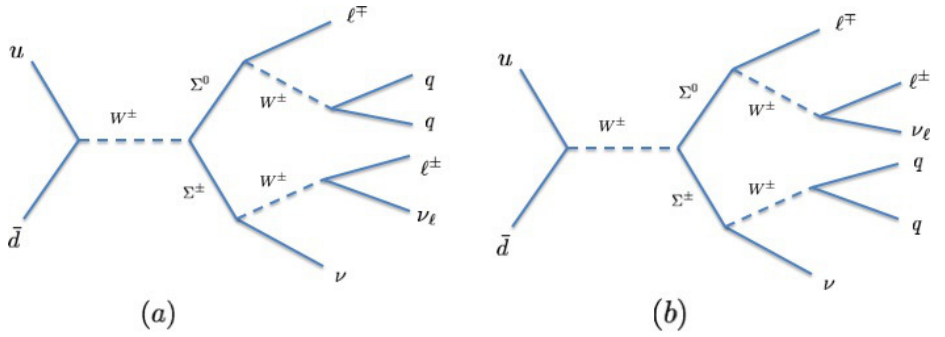


Figure 5.1: Feynman diagram of the process investigated. Diagram (a) will decay with equal likelihood to opposite-sign and same-sign final states. However, diagram (b) will always decay to opposite-sign final states.

the final state is required to have two leptons, and therefore in the case where one or both τ's decay hadronically, the event would be lost.

The requirement of two SS leptons provides a strong discriminant against SM backgrounds. Backgrounds that remain after the SS selection can be separated into three categories based on the origin of the lepton candidates: fake or 'non-prompt' lepton background in which at least one lepton candidate is not a true prompt lepton (*fake leptons*), charge-flipped lepton background in which one lepton is assigned the wrong electric charge (*charge mis-identified leptons*), and true prompt lepton background from SM physics processes involving two SS leptons in the final state. Prompt here means that the leptons originate from the primary vertex, not leptons within showers or other secondary decays.

In this analysis the two electron final state is being studied rather than using muons and τ's as well. However, τ's coming from the heavy neutrino decay could potentially enter into the signal region through its decay to electrons, which happens to $\approx$ 17.8% [18] of all τ's. Leptonically decaying τ's lose energy via two ν's which lowers the efficiency of reconstruction. Only 14% of the SS events used for the cross-section measurement have an electron from a τ decay from the heavy lepton. Hadronic τ decays are not expected to enter

the signal region. The τ's have been constrained to decay only leptonically in the signal samples.

5.2 Monte Carlo Simulation

A common method for simulating the various known Standard Model processes, or exotic undiscovered processes such as the Seesaw Model, is Monte Carlo simulation (MC). Monte Carlo is a probabilistic method which is used to describe the desired process to simulate. It is used to simulate the production and the decay of particles, and their interactions with the detector. Finally, the simulated events are reconstructed in the same manner as the data. Monte Carlo background events are weighted using measured or calculated cross-sections to properly represent the data which is recorded at the LHC. The cross-sections of the most prevalent processes are summarized in Figure 5.2.

The full process of MC generation in summarized in the flowchart in Figure 5.3 [112]. A generator produces events in the HepMC format, which can be filtered at generation for certain requirements. The event information at this stage is also known as truth information, which is the ideal data of a physics process. The generated events are processed through a simulation of the ATLAS detector using GEANT4 [113]. Energies deposited into sub-detector components are recorded as hits, and stored in a simulation output. Digitization takes that hit output and detector noise and pile-up are added. Additionally the L1 trigger is implemented in a "pass" mode, so that no events are discarded, but the trigger value is evaluated. The digitization constructs the inputs to read out drivers (RODs), creating Raw Data Objects(RDOs). The HLT is run on these RDO files. At this stage, reconstruction for data and simulation is identical [112].

Monte Carlo samples often have a 'k-factor' applied to normalize MC events to the number of events measured in data. This k-factor is necessary typically to account for NLO and NNLO QCD corrections which can be difficult to accurately estimate through

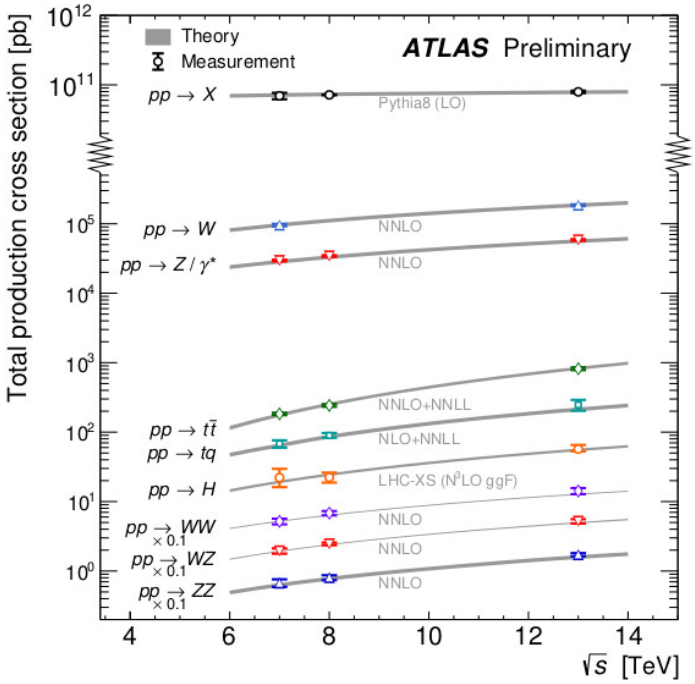


Figure 5.2: Measured and calculated cross-sections of the most common processes at the LHC as a function of center-of-mass energy. The cross-sections also include processes with additional particles. For example $pp \to W$ includes $pp \to W + X$, where X could be additional quarks.

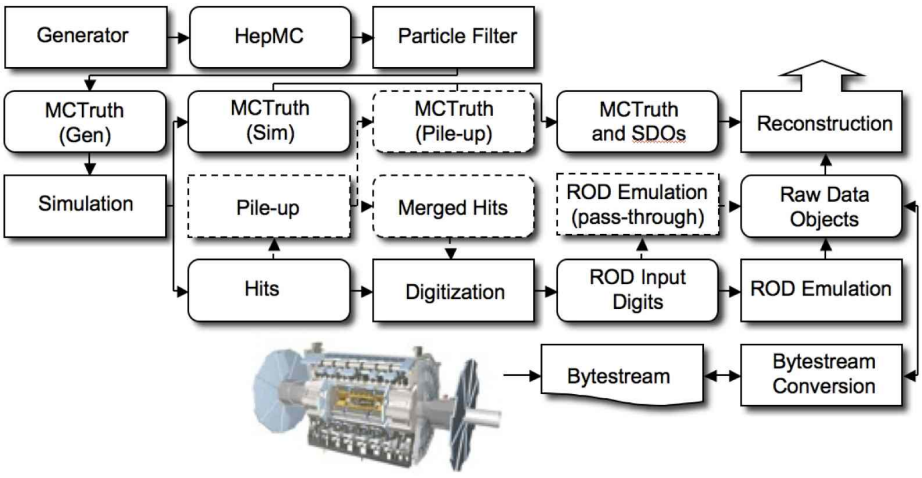


Figure 5.3: Flowchart describing the process of generating Monte Carlo events. Algorithms are represented by square boxes, and stored information is represented by rounded boxes.

simulation. Additionally generator filter previously mentioned also must be applied as a weight on the MC.

The following subsections describe the signal and background processes, simulated with MC, in the Type III Seesaw Heavy Lepton analysis. Backgrounds are considered for one of two reasons; either the final state is similar to the signal final state, or through charge mis-identification or jets faking leptons, the background has a similar signature to the signal final state. Both charge mis-ID and jets faking leptons are very rare processes, so it is only necessary to consider these effects in physical backgrounds with large cross-sections.

5.2.1 Generators

A generator is an algorithm which simulates a specific hard scatter process [114]. The generator starts by simulating the heart of the collision, and calculating from perturbation theory the probability distribution of a particular scatter. There are many generators used at ATLAS with different specializations [115]. Pythia8 [116], POWHEG [117], ALPGEN [118], SHERPA [119], and MadGraph5 [120] are among the most common generators used at ATLAS.

5.2.2 Type III seesaw signal

The signal Monte Carlos were generated with MadGraph [120] and Pythia8 [116]. They were generated at multiple mass points for the heavy leptons: $200, 250, 300, 400, 500, 600$ and 700 GeV.

For the purpose of this analysis, the branching ratios of the heavy lepton decay to electrons, muons, and τ's are all assumed to be equal. It was also assumed that the masses of Σ^0 and $\Sigma^{\pm}$ are identical.

This analysis uses MadGraph5 to create the heavy leptons of the Type III Seesaw model and calculate their decay branching ratios. These signal particles are then used by Pythia8 which simulates their decay. Once all the resultant particles are simulated, GEANT4 [113] is used to simulate their interaction with the ATLAS detector.

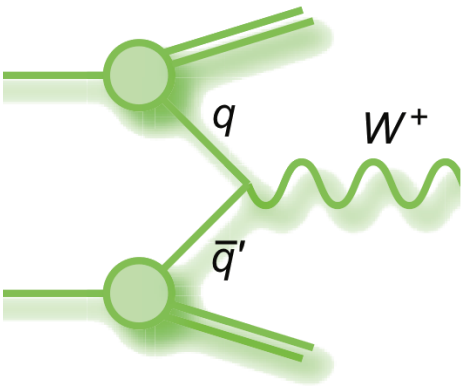
5.2.3 W+Jets

W+jets are a background in this analysis due to their high production rate and non-negligible chance of a jet faking an electron. W bosons are created primarily via quark anti-quark interactions in the pp collisions at the LHC [121]. The leptonically decaying W boson has a cross-section times branching fraction of 11.83 ± 34 nb at 13 TeV at the LHC [122]. With an integrated luminosity of 3.2 fb^{-1}, there are 3.7×10^7 leptonically decaying bosons produced in 2015.

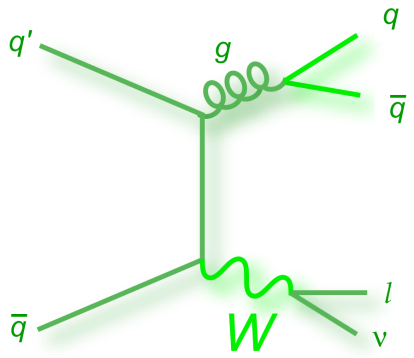
Due to the strong force, Ws are often produced in association with gluons and quarks which produce jets at the LHC. Figure 5.4 shows a W boson produced in association with a gluon which decays to two quarks.

5.2.4 $t\bar{t}$

With a cross-section of 818 ± 36 pb, $t\bar{t}$ pair production is a very common process of pp collisions at the LHC [123]. The signature in the detector is very similar to the signal of this analysis. As mentioned in Chapter 2, the top quark decays instantly, and very nearly



(a) Dominant W boson production mode at the LHC



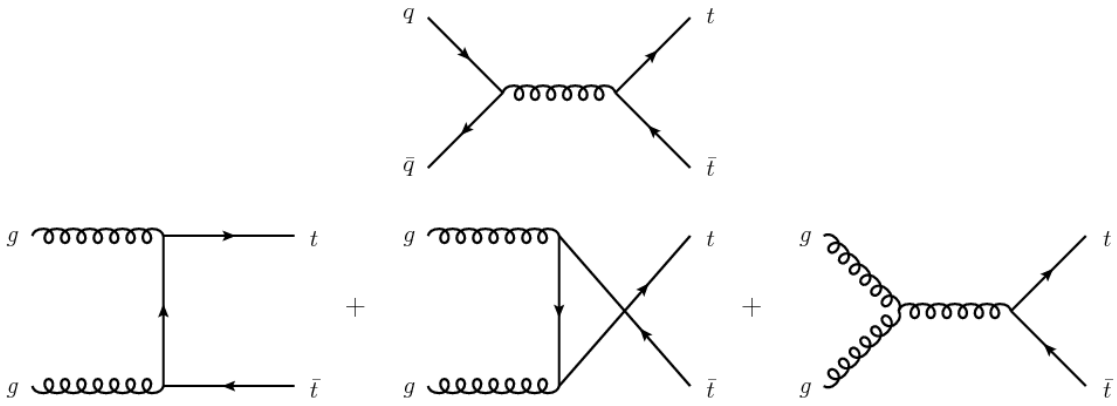
(b) W boson production in association with 2 jets

Figure 5.4: W boson production.

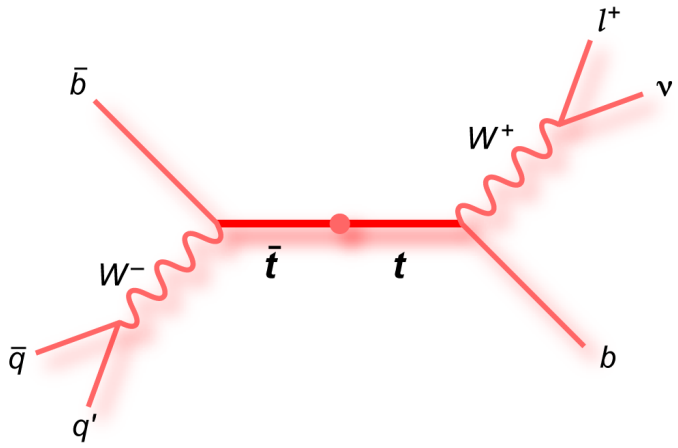
exclusively to a W boson and a b quark. When both the Ws from the top quarks decay to electrons the signature looks very similar to the Type III Seesaw signal. However, in the case of same-sign decays, $t\bar{t}$ contributes to the background when one of the electron's charges is mis-identified. $t\bar{t}$ can also contribute when it undergoes semi-leptonic decay, see Figure 5.5b, and one of the jets fakes an electron signature. The only difference is that there are four expected jets in the detector, and two of these jets come from a b-quark. b-tagging, described in Section 4.5, can be used to tag these jets. Then events with b-tagged jets can be rejected since b-jets are not expected in the signal process.

5.2.5 Single Top Quark

Single top quark production occurs at a cross-section of 156 ± 28 pb at the LHC [126]. Single top quarks can be produced in several channels, the most common of which is shown in Figure 5.6. It is characterized by at least one b-jets in the final state. Single top processes can only contribute to the signal region when one of the jets fakes a lepton and when the W boson produced decays leptonically. Consequently this background contributes the fewest events in the signal region where the signal cross-section measurement is made.



(a) Lowest order Feynman diagrams of $t\bar{t}$ pair production [124].



(b) Semi-leptonic $t\bar{t}$ decay [125].

Figure 5.5: $t\bar{t}$ production mechanisms at the LHC and $t\bar{t}$ semi-leptonic decay.

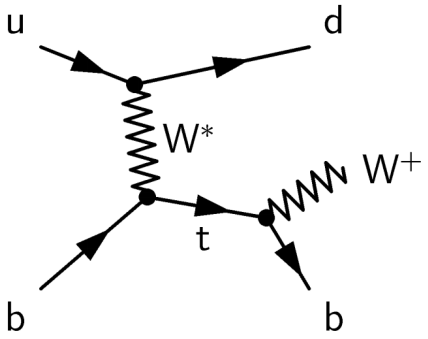


Figure 5.6: Leading order single top production mechanism [126].

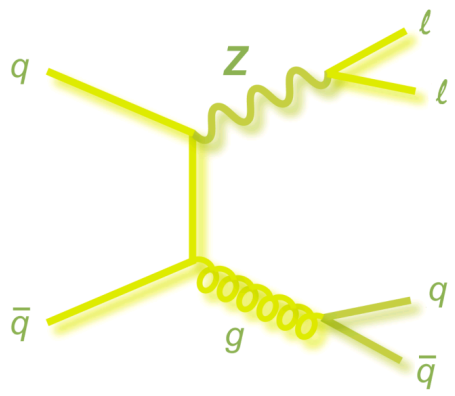


Figure 5.7: Z boson production with two jets.

5.2.6 Z+Jets

The Z boson decaying to two electrons has a cross-section times branching fraction of 778 ± 17 pb at 13 TeV at the LHC [127]. With a luminosity of 3.2 fb^{-1}, there are nearly 2,500,000 e^+e^- pairs from Z bosons produced in 2015.

Z bosons can contribute to the signal region when they radiate two or more jets in the event, see in Figure 5.7. These events always have two opposite-sign leptons, however, because of the high rate of production of the Z boson at the LHC, Z bosons can be a background in the analysis through charge mis-identification. For electrons with a p_{T} around 45 GeV the average charge mis-identification rate is $\approx 1.6\%$, meaning a potential 250,000 same-sign events. Charge mis-identification is described in more detail in Section 5.4.2.

5.2.7 Diboson

The diboson contribution is one of two physical background contributions with two true prompt electrons in the same-sign channel. WZ boson production has a cross-section of 50.6 ± 3.6 pb [128] and can produce a final state with three leptons. It is possible that one of those leptons is not reconstructed in the detector, resulting in a same-sign final state. For example, W^+Z can decay leptonically to $e^+\nu_e e^+e^-$. If the electron passes into a region

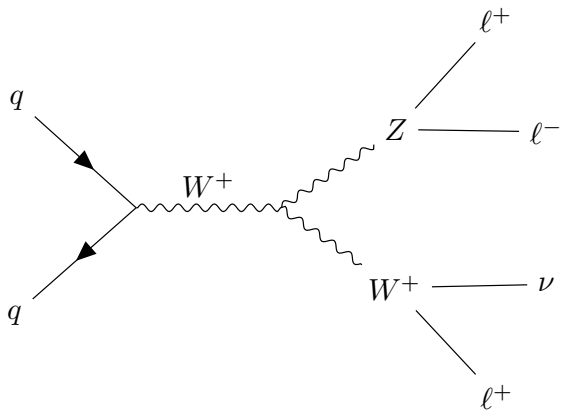


Figure 5.8: Diboson production with leptonically decaying W and Z bosons.

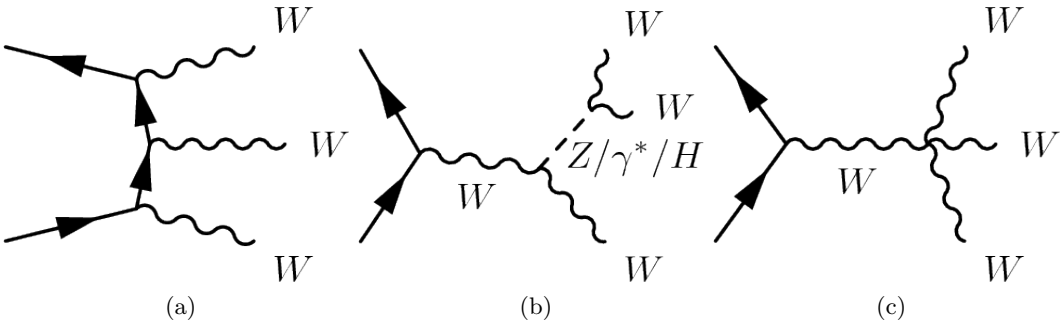
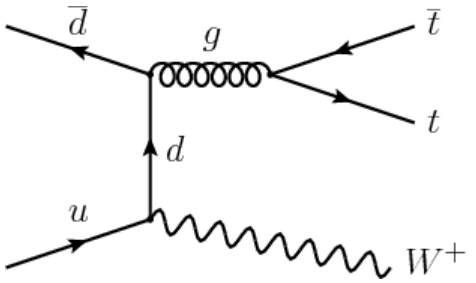


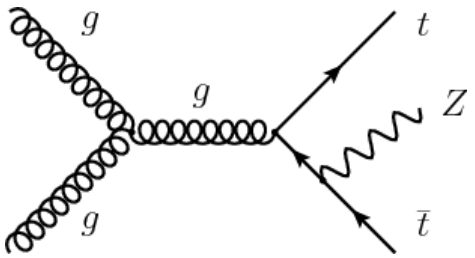
Figure 5.9: Leading order triboson production

of the detector where it cannot be reconstructed, there are two positrons and some missing energy in the final state, see Figure 5.8.

Furthermore, triboson production, for example $W^+W^+W^-$, can directly create a same-sign final state that is identical to the Type III Seesaw signal in this analysis: $e^+\nu^e e^+\nu^e + 2$ jets. Figure 5.9 shows the most common production mechanisms for WWW [129]. The cross-section for this process is 241.5 fb^{-1} [129]. Specifying the same-sign electron and two jets final state results in a cross-section times branching fraction of 5.85 fb^{-1}.



(a) $t\bar{t}$ produced in association with a W boson



(b) $t\bar{t}$ with a Z boson radiated from one of the final state top quarks

Figure 5.10: $t\bar{t}+V$ production [131].

5.2.8 $t\bar{t}+V$

$t\bar{t}+V$ is a $t\bar{t}$ decay where a W or Z boson is also produced. The boson can be a W or Z boson, Figure 5.10. $t\bar{t}+Z$ has a cross-section of 0.9 ± 0.3 pb, and $t\bar{t}+W$ has a cross-section of 1.4 ± 0.8 pb at 13 TeV at the LHC [130]. Both of these processes can produce same-sign electrons with multiple jets and $E_{\mathrm{T}}^{\mathrm{miss}}$. They can be very difficult to separate from signal, however, b-jet tagging and a relatively low cross-section makes $t\bar{t}+V$ a minor background in this analysis.

5.3 Charge Mis-identified Electrons

e^+e^- pairs are produced at a high rate in pp collisions, via Z bosons or photons [132]. Charge mis-identification is a rare occurrence, however due to the quantity of e^+e^- pairs being produced, it is one of the dominant backgrounds in this analysis. Charge mis-identified events are generated implicitly via the GEANT4 simulation of the ATLAS detector [113]. However the rate for charge mis-identified events in GEANT4 simulations does not perfectly match with the results collected in data. The Monte Carlo estimation needs an additional correction in order to achieve good agreement with data.

One alternative to using the MC generated same-sign events is to calculate the charge flip rate and apply this to the opposite-sign data events, as described in Section 5.4.2.

5.4 Data Driven Background Estimation

Monte Carlo can be produced at very high statistics and weighted to model the physical processes recorded in data.

The high statistics available through Monte Carlo is advantageous for shape studies and machine learning. Large numbers of data points allows machine learning to be trained more sophisticatedly. Furthermore, the data has smoother shapes, allowing more accurate shape fitting.

Weighting is done by applying a value, event by event, to increase or decrease a particular event's contribution to the total estimation of any distribution. Weighting is performed for two main types of corrections: data-MC discrepancy corrections and scale factors to match data luminosity. In this section, weighting is used to estimate a background using data which has similar kinematic properties, but different rates.

One of the goals of this analysis was to study if multivariate machine learning could be used to better separate signal from background than simpler cut based methods. In order to use the multivariate method, large samples of both background and signal are crucial. However, when examining same-sign final states, the dominant backgrounds are due to fake leptons and charge mis-identification. These processes are not available in Monte Carlo with large numbers of events suitable for machine learning methods.

An alternative to MC modeling is to use data driven approaches, which estimate contributions due to mis-reconstruction directly. This method allows much higher statistics in the signal region. Monte Carlo subtraction using truth information must be done to avoid double counting contributions from fake electrons and charge mis-identification. For example, the data-driven charge mis-identification background includes the Z boson's contribution. Using the Z+jets Monte Carlo would double the contribution of the Z+jets as a background. After subtracting events by using truth information, only the diboson and $t\bar{t}+V$ Monte Carlo still contribute to the final state.

5.4.1 Fake Electrons

A source of significant background in same-sign events can come from the mis-identification of jets in the detector, and from non-prompt leptons. These objects can be mis-identified as electrons. The fake signals come from three main sources: light flavor jets, heavy flavor jets, and photon conversions [133].

Jets can occasionally produce a large fraction of their energy in the EM calorimeters compared to the TileCal, and this may be reconstructed as an electron. Heavy flavor jets may produce non-prompt electrons in semi-leptonic c or b quark decays. Typically these electrons will fail the isolation requirements, however, occasionally the kinematics of a decay will allow the non-prompt electrons to pass the isolation cuts.

Photons, either from the hard scatter of the event, from the decay of another particle, or produced from bremsstrahlung, may produce electrons through pair production. The photon can produce an electron-positron pair and one or both of the resulting particles can be reconstructed.

It is possible to use models to simulate each of these signatures individually, however they have large uncertainties and it is quite difficult to validate the agreement between data and simulation. In light of these difficulties, the alternative is to group these processes together using a data-driven approach which does not distinguish the sources of the fake electrons. In this analysis a fake factor method [134] is employed to estimate the backgrounds due to the fake electrons.

5.4.1.1 Electron Fake Rate

This analysis uses a data-driven method to determine the electron fake rate estimation described in [135]. In the fake factor method, objects satisfying all of the loose electron selection criteria are referred to as denominator (D) objects. A sub-sample of this set that additionally satisfy the tight selection are referred to as numerator (N) objects. Ideally both the sample for calculating the fake rate and sample used for the estimation are composed

entirely of fake electrons. The ratio between N and D defines the fake factor f. This rate is essentially the efficiency of loose non-electron objects to be identified as a tight electron.

Ideally the electron fake rate would be measured with a data sample in which all the electrons are fake. However, in practice it is difficult to make such a selection with sufficient fake electron purity and high efficiency. Instead a fake electron enriched region is created, from which the remaining prompt electrons are subtracted using MC.

The fake electron enriched region is created from a set of single high-p_{T}, loose electron triggers listed in Table 5.1. These are electron triggers similar to those described in Section 4.3.1. These triggers differ from the previous by being looser and less stringent in their identification of electrons.

Several cuts are imposed to suppress events with prompt electrons. These cuts define the fake enriched region. Events must have at least one loose electron and pass the following requirements:

- $Z \to ee$ veto: reject events with $m(ee)$ close to the Z mass ($71.2 < m(ee) < 111.2$ GeV)
- Drell-Yan [132] veto: reject events with two or more tight electrons
- $W \to e\nu$ veto: reject events with $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV

This selection primarily selects di-jet events. The p_{T} distributions for the selected tight and loose electrons in the fake electron enriched region are shown in Figure 5.11. The figure demonstrates that the majority of events in this fake enriched region are not due to real electrons, which are modelled by the MC backgrounds. The number of events with real electrons with the loose requirement are not significantly diminished when requiring tight requirements. However for the data, which mostly contains fake electrons, there is a decrease by nearly a factor of 5.

The selected electrons in the fake enriched region are split into nine p_{T} bins and three η bins. In each of the 27 η-p_{T} bins the number of loose (tight) electrons is calculated as

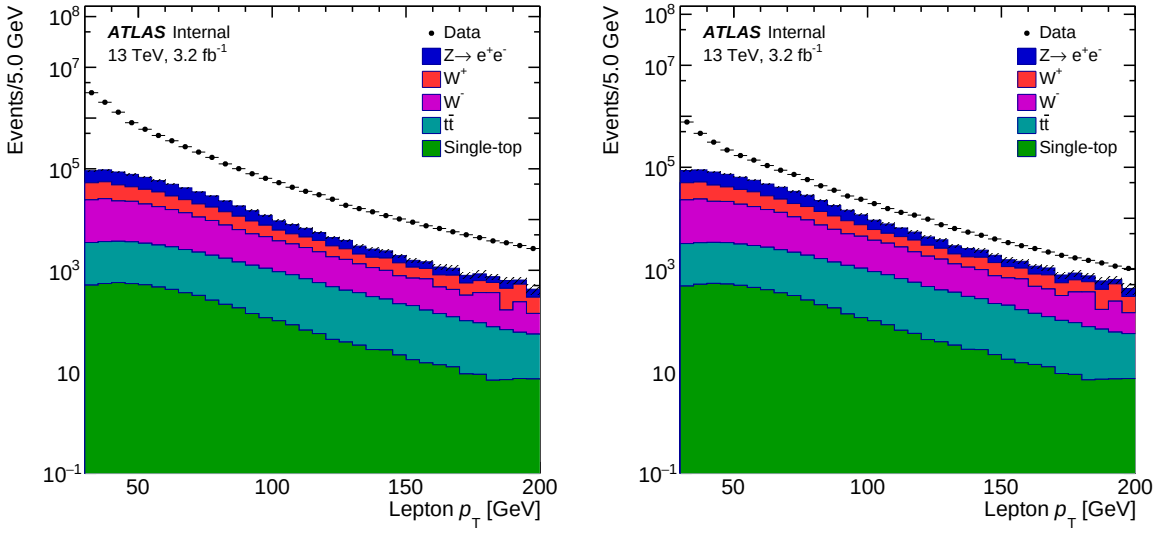


Figure 5.11: Electron p_{T} distributions in fake electron enriched region for data and prompt electron MC samples for (a) loose electrons (D), (b) tight electrons (N). The prompt electron backgrounds shown are from MC simulation. The excess data above the MC background is due to fake electrons in the data.

the raw yield of loose (tight) electrons minus the prompt electron yield estimated from MC simulated events (W+jets, Drell-Yan[†], Single top quark, and $t\bar{t}$). The fake electron rates for the η-p_{T} bins are shown in Figure 5.12b and listed in Table 5.2. As the analysis final state has at least 2 jets, the fake rate used in the analysis was determined with the same requirement.

Table 5.1: Set of single electron triggers used for the fake electron enriched region.

Trigger	p_{T} range [GeV]	Trigger prescale
HLT_e24_lhvloose	30 – 65	35
HLT_e60_lhvloose	65 – 125	12
HLT_e12_lhloose	125 – ∞	1

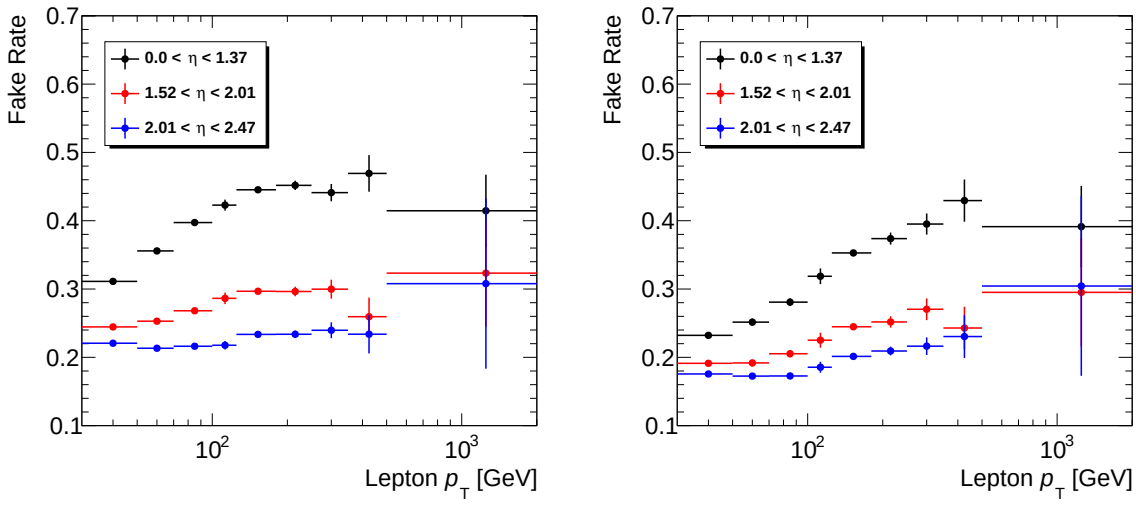
[†]The Drell-Yan contributions are included in the $Z \to e^+e^-$ sample for all figures

Table 5.2: Electron fake rate for events with ≥ 2 jets in $\eta - p_{\mathrm{T}}$ bins determined in the fake electron enriched region.

p_{T} [GeV]	$0 < \eta < 1.37$	$1.52 < \eta < 2.01$	$2.01 < \eta < 2.47$
$30 < p_{\mathrm{T}} < 50$	0.232 ± 0.030	0.191 ± 0.021	0.176 ± 0.021
$50 < p_{\mathrm{T}} < 70$	0.252 ± 0.060	0.192 ± 0.029	0.173 ± 0.021
$70 < p_{\mathrm{T}} < 100$	0.281 ± 0.040	0.205 ± 0.024	0.173 ± 0.018
$100 < p_{\mathrm{T}} < 125$	0.319 ± 0.025	0.225 ± 0.014	0.186 ± 0.012
$125 < p_{\mathrm{T}} < 180$	0.353 ± 0.021	0.245 ± 0.010	0.201 ± 0.006
$180 < p_{\mathrm{T}} < 250$	0.374 ± 0.018	0.252 ± 0.015	0.209 ± 0.015
$250 < p_{\mathrm{T}} < 350$	0.395 ± 0.024	0.270 ± 0.031	0.216 ± 0.017
$350 < p_{\mathrm{T}} < 500$	0.429 ± 0.034	0.243 ± 0.032	0.230 ± 0.033
$500 < p_{\mathrm{T}}$	0.391 ± 0.087	0.295 ± 0.079	0.304 ± 0.132

Systematic uncertainties in the electron fake rate arise from MC modeling and normalization of the MC estimate of prompt electrons in the fake enriched region and from different compositions of fake and non-prompt electrons in the fake enriched region from the signal region.

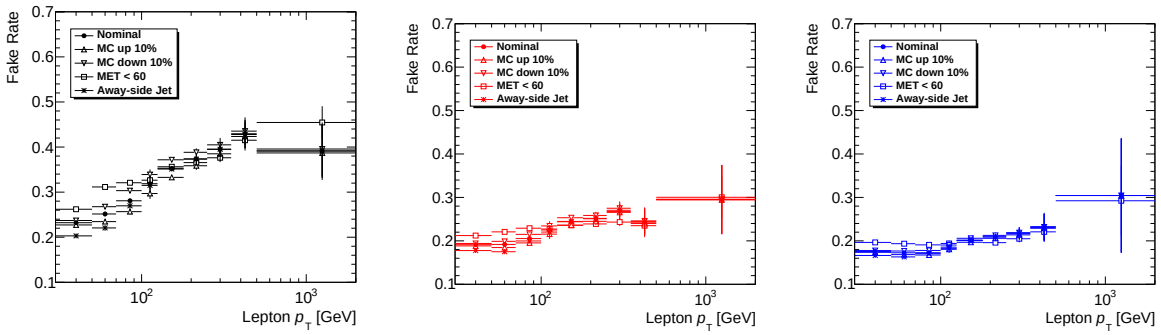
The systematic uncertainties in the electron fake rate are estimated by applying variations to the nominal selection of the fake enriched region. The largest contribution of prompt electrons in the fake electron enriched region comes from W+jets events. In order to test the modeling of the W+jets MC simulated events the nominal $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement of > 25 GeV is loosened to 60 GeV. The normalizations for all MC samples have been varied by 10% up and down to allow for uncertainties in the cross-sections. An away-side jet (i.e. a jet with $\Delta\phi$(jet, electron) > 2.4) is required for each electron in the event in order to tighten the specification of di-jet processes in the fake enriched region. The estimated fake rate for each variation are shown in Figure 5.13. The total systematic uncertainty is calculated by adding all of the variations in quadrature. The final fake rates including systematic and statistical uncertainties are shown in Figure 5.14.



(a) p_{T} dependence of the electron fake rate in the selected events

(b) p_{T} dependence of the electron fake rate in the selected events with ≥ 2 jets

Figure 5.12: p_{T} dependence of the electron fake rate. These rates are determined from the fake enriched regions after subtraction of the prompt electron background.

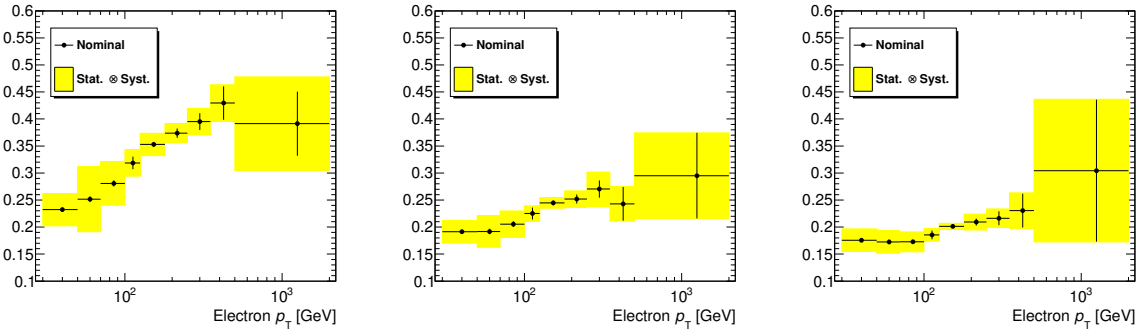


(a) $0 < \eta < 1.37$

(b) $1.52 < \eta < 2.01$

(c) $2.01 < \eta < 2.47$

Figure 5.13: Systematic variations versus fake electron p_{T} for events with 2 or more jets.



(a) $0 < \eta < 1.37$ (b) $1.52 < \eta < 2.01$ (c) $2.01 < \eta < 2.47$

Figure 5.14: p_{T} dependence of the electron fake rate for 2 or more jets including systematic and statistical uncertainties.

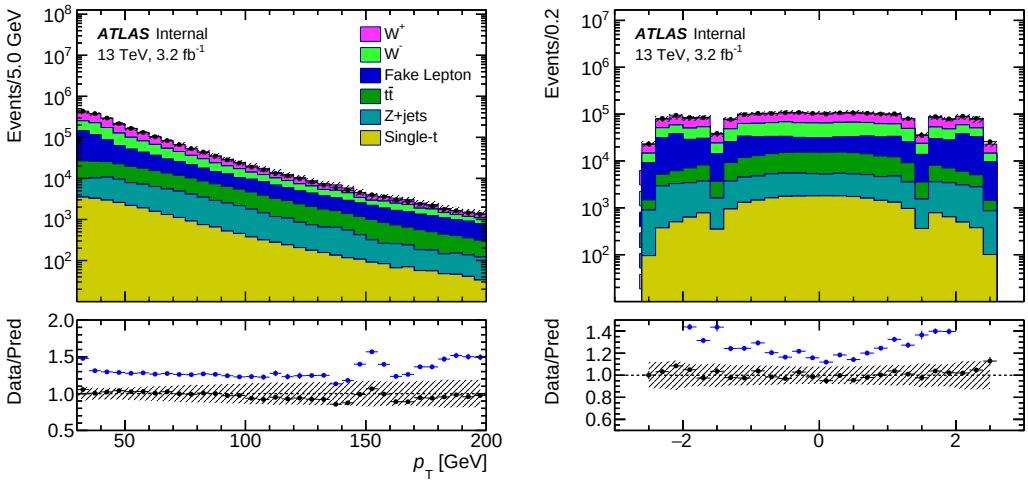
A fake lepton sample was generated for the W+jets enriched region. This region is defined by requiring the following:

- Exactly 1 tight electron
- $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV
- $M_{\mathrm{T}} > 50$ GeV
- At least 2 jets

M_{T} is the transverse mass, defined as a relation between the energy and momentum of the neutrino and lepton in the event [136]:

$$M_{\mathrm{T}}^2 = 2E_{\mathrm{T}\nu}E_{\mathrm{T}l} - 2\vec{p}_{\mathrm{T}\nu} \cdot \vec{p}_{\mathrm{T}l}$$

It is expected that a fairly large fraction of events in the W+jets enriched region are actually due to fake electrons. Thus the events measured in data are not accounted for by the prompt electron MC alone. To account for the fake leptons, the loose electron sample is used to estimate the fake electron contribution in the W+jets region which requires a tight electron. The fake rate is used as a weight, dependent on the electron's η and p_{T}. More details about how this sample is estimated are given in Section 5.4.1.2.



(a) p_{T} distribution of highest p_{T} electron

(b) η distribution of highest p_{T} electron

Figure 5.15: W+jets validation region for the electron fake rate. All events are required to have at least 2 jets. The blue points show the Data-Background comparison without the Fake Lepton contribution, whereas the black points include Fake Lepton contributions.

Figure 5.15 shows the comparison of the p_{T} and η distributions between data and MC for the highest p_{T} electron (leading electron) in a given event. The blue points in the lower plots represent the ratio between data and MC without accounting for fake electrons, while the black points also consider the fake electron estimate as part of the background. The Monte Carlo normalization used for this particular comparison does not agree well with data with two or more jets in the event. This is due to a known issue with the POWHEG generator used to produce the W boson samples. To account for this issue, the normalization of the W boson contributions have been scaled up by $\approx 35\%$. This factor was estimated with a best fit approach using the M_{T} distribution. Further justification for the normalization factor is detailed in Appendix A.

5.4.1.2 Generating a Fake Electron Sample

In an event with two electrons, the fake factor method has to account for each combination of fake and real electrons. Either of the two electrons $\ell_1\ell_2$, could be either real (ℓ_i^{R}) or

fake (ℓ_i^{F}). The fake factor method provides a way of estimating the number of $\ell_1^{\mathrm{R}}\ell_2^{\mathrm{R}}$ events in a sample of $\ell_1^{\mathrm{N}}\ell_2^{\mathrm{N}}$ observed candidates in data.

Since either electron can be fake, the observed dielectron events in the analysis can be decomposed as:

$$\ell_1^{\mathrm{N}}\ell_2^{\mathrm{N}} = \ell_1^{\mathrm{R}}\ell_2^{\mathrm{R}} + \ell_1^{\mathrm{R}}\ell_2^{\mathrm{F}} + \ell_1^{\mathrm{F}}\ell_2^{\mathrm{R}} + \ell_1^{\mathrm{F}}\ell_2^{\mathrm{F}}, \tag{5.1}$$

where ℓ_1 and ℓ_2 denote the first and second leading electron in the event. Now, for the case where one of the electrons is loose, which is the electron selection used to generate the complementary sample (D), the relation becomes:

$$\ell_1^{\mathrm{D}}\ell_2^{\mathrm{N}} f_1 = \ell_1^{\mathrm{F}}\ell_2^{\mathrm{R}} + \ell_1^{\mathrm{F}}\ell_2^{\mathrm{F}}, \tag{5.2}$$

$$\ell_1^{\mathrm{N}}\ell_2^{\mathrm{D}} f_2 = \ell_1^{\mathrm{R}}\ell_2^{\mathrm{F}} + \ell_1^{\mathrm{F}}\ell_2^{\mathrm{F}}, \tag{5.3}$$

where ℓ_i^{D} denote electrons with the loose selection, and f is the fake factor. Similarly, where ℓ_1 *and* ℓ_2 are both loose:

$$\ell_1^{\mathrm{F}}\ell_2^{\mathrm{F}} = \ell_1^{\mathrm{D}}\ell_2^{\mathrm{D}} f_1 f_2. \tag{5.4}$$

Substituting Equations (5.2–5.4) into Equation (5.1) gives:

$$\ell_1^{\mathrm{N}}\ell_2^{\mathrm{N}} = \ell_1^{\mathrm{R}}\ell_2^{\mathrm{R}} + (\ell_1^{\mathrm{D}}\ell_2^{\mathrm{N}} f_1 + \ell_1^{\mathrm{N}}\ell_2^{\mathrm{D}} f_2 - \ell_1^{\mathrm{D}}\ell_2^{\mathrm{D}} f_1 f_2), \tag{5.5}$$

where the total background due to the fake electrons is the one inside the bracket.

In order to generate a sample to estimate the fake electron contribution in the signal region, the fake rate is applied to the strictly loose electrons which meet all the requirements of the signal region excluding the tight electron definition. The fake rate is used to find a ratio applicable in the strictly loose region by using $F = 1/(1-f)$ and taking F to be the ratio of strictly loose electrons to tight fake electrons. The ratio F is applied as a weight on each event, depending on the strictly loose electrons' p_{T} and η. If both electrons are loose, then the weight is multiplied by -1, consistent with Equation 5.5.

However, these strictly loose same-sign electrons may actually be opposite-sign electrons where one of the electron's charges is mis-identified, the charge mis-identified background is described in Section 5.4.2. Additionally, there will also be some contribution from true same-sign processes which can show up due to diboson processes or through $t\bar{t}$ processes in association with a vector boson. The charge mis-identified events are subtracted as well as true SS processes before applying the fake factor.

The same-sign fake electron estimation calculation is summarized in the equation below:

$$\text{Fake contribution} = f_{\text{rate}} \times (\text{SS loose electrons from jets})$$

To select the loose electrons originating from jets specifically, it is necessary to subtract from the SS loose objects, the Charge Mis-ID'd objects and the true, prompt SS processes.

$$\text{Fake contribution} = f_{\text{rate}} \times (\text{SS} - \text{OS} * \text{MisID}_{\text{rate}} - \text{MC}_{\text{prompt}})$$

Once these corrections are made, the same-sign fake electron contribution is well modeled and agrees well with data. However the number of events with same-sign electrons is small. In order to have a larger data sample with smaller relative statistical uncertainty for the analysis, events with same-sign and opposite-sign electrons are used and normalized to the number of events with same-sign electrons. Further motivation for the normalization is detailed in Appendix B.

The process to estimate same-sign fake electron contribution is summarized below:

The signal region is specified as described earlier, with the exception that at least one of the electrons in the event must fail the tight electron definition but pass the loose. The electrons are also not required to be same-sign. This sample of events is mutually exclusive with those in the signal region. For each event, a scale factor is calculated based on the electron's η and p_{T}. Similarly, scale factors are calculated for MC which is used to remove the prompt electrons. The charge mis-identified sample is subtracted and the resulting sample is then normalized to the same-sign only contribution.

Table 5.3 shows the contribution of same-sign fake electrons and the other contributions in the SS loose selection that have been subtracted.

Table 5.3: Summary of the sources of events which contribute to the SS region. SS contributions from charge Mis-ID sources and prompt sources are subtracted

Source	Count
All SS objects	28.0
SS from charge Mis-ID	4.1
SS from prompt	0.5
SS from fakes alone	23.4

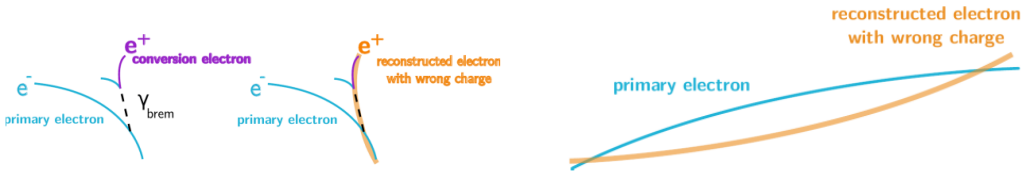
5.4.2 Charge Mis-identified Electrons

Another important source of background are events with from opposite-sign true electrons, in which the charge of one electron is mis-reconstructed. An electron's charge can be mis-identified through two main processes.

The first process consists of an electron producing a bremsstrahlung photon which subsequently converts to an electron positron pair. The reconstruction algorithm incorrectly identifies the wrong-charge secondary electron as the prompt electron, while losing the primary electron. This process is shown in Figure 5.16a. A second possibility, shown in Figure 5.16b, is when the electron has a very high p_{T}, such that the electron trajectory in the detector is almost straight. As the reconstruction algorithm uses the sign of the track curvature to assign a charge to the electron, the assigned charge has a high probability to be mis-reconstructed.

5.4.2.1 Charge Flip Rate Measurement

A data-driven approach is used to estimate the contribution of electron charge misidentification within the signal region [137]. Events in which a Z boson decays to an e^+e^- pair are utilized as a source of background free opposite-sign electrons. Near the Z mass peak, it is expected that the majority of electrons are due to the Z boson decaying to



(a) Conversion of a bremsstrahlung photon. (b) Wrong track reconstruction.

Figure 5.16: The two dominant mechanisms by which electron charge is mis-measured.

electron-positron pairs. Due to the high cross-section of the Z boson, virtually all of the same-sign electron pairs near the Z peak result from charge mis-identification of one of the electrons.

However, there are other sources of electron pairs besides the Z boson in the electron pair channel. Two regions are defined for the same-sign and any-sign selections: the Z mass region and the Z mass sidebands. These regions are defined in Table 5.4. The same-sign region is shifted by 2 GeV to account for the average lost energy in the charge flip process. Figure 5.17 shows the main regions for the any-sign and same-sign selections. This measurement was performed by another group which utilized the 2015 and 2016 data collected by ATLAS. However, it will be shown that the charge flip rates from this dataset are valid for the 2015 dataset.

Table 5.4: Definition of the Z mass peak and Z mass sideband regions used for the charge mis-ID rate estimation.

Type	Z mass peak	Z mass sidebands
Any-Sign	$\lvert m(ee) - m(Z)\rvert < 14$ GeV	14 GeV $< \lvert m(ee) - m(Z)\rvert < 28$ GeV
Same-Sign	$\lvert m(ee) - m(Z) + 2\rvert < 15.8$ GeV	15.8 GeV $< \lvert m(ee) - m(Z) + 2\rvert < 31.6$ GeV

The sidebands are used to subtract events from non-Z backgrounds from the Z mass peak so that the remaining same-sign contribution is only due to mis-identified electrons

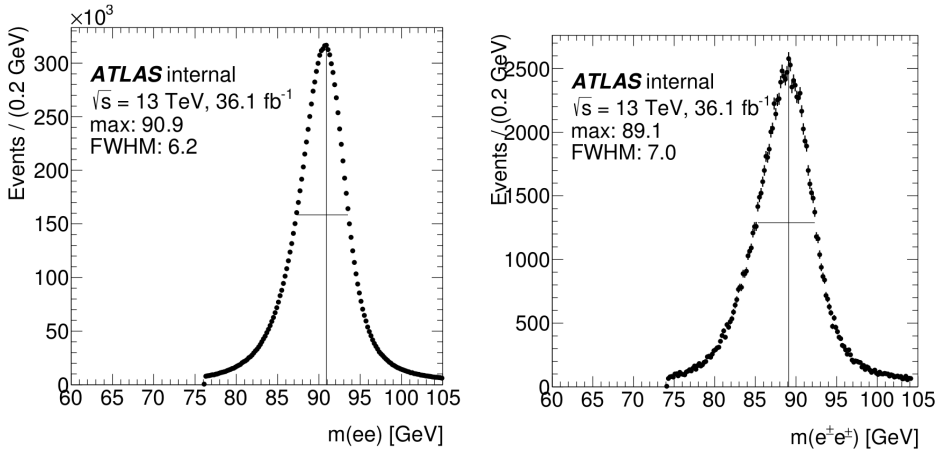


Figure 5.17: Lepton pair mass for $Z \to ee$ (a) without lepton charge requirements; therefore mostly OS, and (b) for the same-sign lepton pair only. There is a shift in the mass peak position by about 2 GeV. The shift is described later in the text [137].

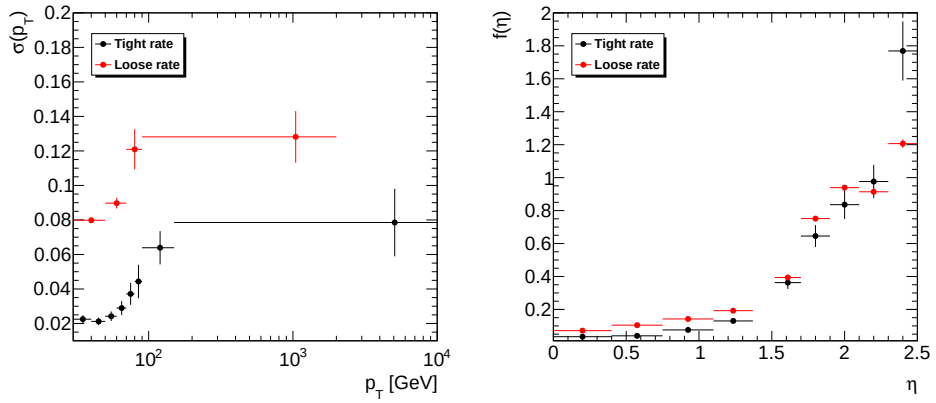
from the Z decay. The sideband method is effective as the non-Z processes form a flat background across the main and sideband regions [135].

A Poisson distribution is commonly used to model the number of times an event occurs in a given space. The same-sign and any-sign events are split into $\eta - p_{\mathrm{T}}$ bins and modeled with a Poisson distribution. Equation 5.6 gives the Poisson probability of observing a same-sign event N_{SS}^{ij}, given the expected number of same-sign events in each i,j bin, where i, and j are each two dimensional bins of η, and p_{T}. The i bin represents the electron with the largest momentum in the events, and the j bin represents the second highest momentum electron.

$$f(N_{SS}^{ij}, \lambda) = \frac{\lambda^{N_{SS}^{ij}} e^{-\lambda}}{N_{SS}^{ij}!} \tag{5.6}$$

Where λ is the expected number of same-sign events in bin i,j:

$$\lambda = (\epsilon_i(1-\epsilon_j) + \epsilon_j(1-\epsilon_j))N_{SS}^{ij}$$



(a) p_{T} dependence of charge mis-ID rate for electrons

(b) η dependence of charge mis-ID rate for electrons

Figure 5.18: Electron charge mis-ID rate versus p_{T} and η.

ϵ_i, ϵ_j are the charge mis-identification probabilities for the two electrons, and N^{ij} is the measured number of events in bin (i,j). The charge mis-identification rates are extracted by constructing a likelihood function from Equation 5.6 and finding the minimum by varying the charge flip probabilities. The total charge flip probability is then given by the η and p_{T} rates, such that the likelihood that an electron's charge has been mis-identified is given by:

$$p_{\mathrm{flip}} = f(\eta) \times \sigma(p_{\mathrm{T}})$$

Figure 5.18 shows the charge mis-ID rate with respect to a given electron's η and p_{T}. Figure 5.19 shows the η dependent charge mis-ID rate split into p_{T} slices, to show that the influence of the variables is uncorrelated.

The systematic uncertainties in the charge mis-ID rate originate from two main sources:

- Fit uncertainty of the likelihood fit parameters
- Selection of the Z mass regions and side-bands

Each of these uncertainties are evaluated in order to provide a symmetric uncertainty of the charge flip method. The fit uncertainties are obtained directly from the fit calculation.

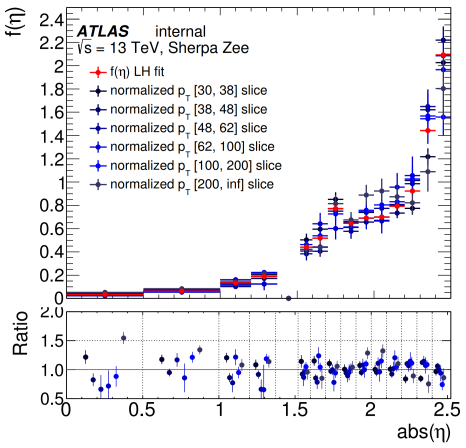


Figure 5.19: Electron charge mis-ID rate dependence on η. The rate was examined within p_{T} slices [137].

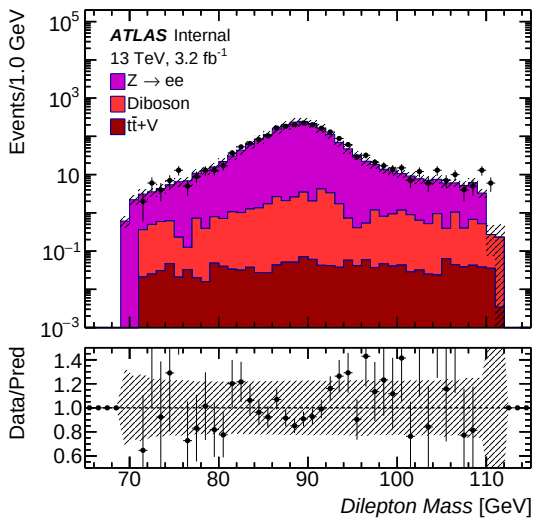
The systematic variation on the selection of the Z mass window is assessed by varying the position of the Z mass window and sidebands by ± 2 GeV. The combination of these two sources of uncertainty gives the total uncertainty of the charge flip method.

The charge mis-ID rate is used as a weight (W) on opposite-sign events depending on the η and p_{T} of the two electrons in the event. W is defined according to Equation 5.7, where p_1 is the charge mis-ID rate of the highest p_{T} electron in the event and p_2 is the charge mis-ID rate of the second electron.

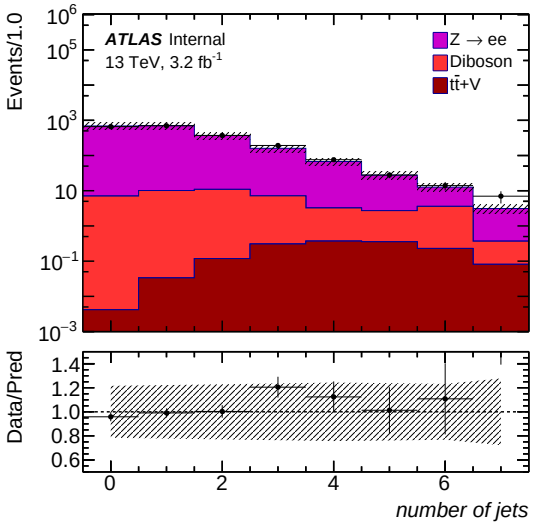
$$W = p_1 + p_2 - 2p_1p_2 \tag{5.7}$$

The distributions in Figure 5.20 show good agreement between the data and predictions in the Z boson mass peak region. The diboson and $t\bar{t} + \mathrm{V}$ backgrounds are estimated via MC, while the $Z \to ee$ is estimated using the data-driven method described. Figure 5.20a shows that the invariant mass peak shifts 2 GeV lower in the data than the prediction. The shift results from the loss of energy due to bremsstrahlung. This Z mass peak region is excluded from the signal region, as the fake electron estimation is not valid here. In this

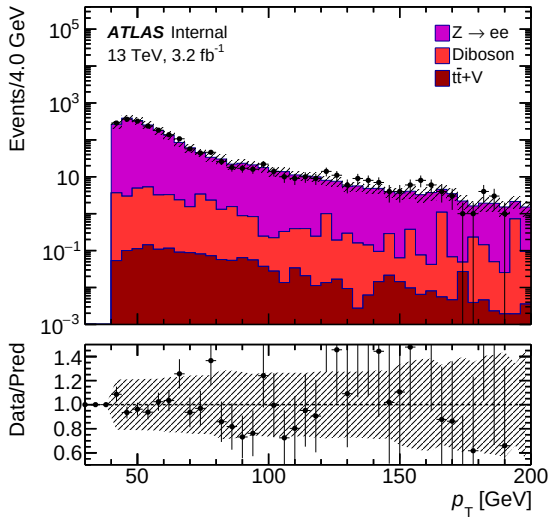
region the majority of loose electrons are real and not due to jets, and therefor the loose electrons of this region fail the assumption required for the fake electron estimation.



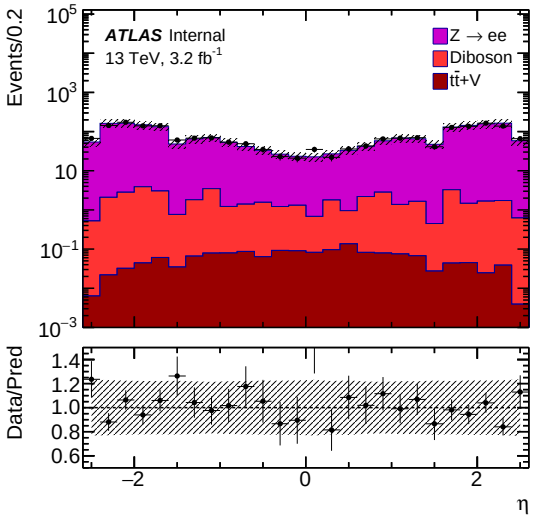
(a) $Z \to ee$ mass with the charge mis-ID contribution shifted by -1.5 GeV



(b) Jet multiplicity



(c) Leading electron p_{T}



(d) Leading electron η

Figure 5.20: Plots of kinematic variables for SS selected events in the Z boson mass region. The hatched areas represent the combined systematic and statistical errors.

CHAPTER 6. ANALYSIS DESCRIPTION

In this chapter, the details of the analysis are described. Control regions are described to validate that there is good agreement between data and the background estimates.

The goal of this analysis is to search for the Type III Seesaw heavy neutrino by looking for an excess of events in a region where signal over background is high, the signal region. If no excess is found, an exclusion interval is calculated to place an upper limit on the possible cross-section of the signal. Two methods to enhance the signal to background ratio are compared. Multivariate techniques are tested and compared to the cut optimized method previously used in the Run 1 analysis.

This chapter covers the selection of the signal region which is treated in a blinded manner. This means that when examining the signal region the data is excluded. In this way there is no experimental bias in the choices of kinematic cuts or multi-variate variable selection. After a few cleaning cuts are performed to ensure background estimations are well modeled, kinematic variables are used to train several methods of multi-variate analysis (MVA). Using the output discriminants produced by the MVA, further selection of the signal region is performed. After the analysis selection is finalized and the measurement sensitivity is calculated, the analysis is unblinded and the data examined.

6.1 Background Control Regions

Control regions are kinematic regions outside of the signal region. Typically, control regions are chosen which are dominated by one well understood process. Control regions are examined to have control over specific background processes in the signal region. Additionally, there are corrections that are crucial to apply depending on the samples or the triggers used. These correction factors must be checked by using well understood interac-

tions. Regions such as the Z+jets enriched region are useful as the Z+jets production is well understood.

6.1.1 Pile-up Corrections

Accurate modeling of the low p_{T} processes can be very difficult. As a result, the distribution of the average number of proton-proton interactions per bunch crossing does not match well between Monte Carlo and data. The distribution of the average number of interactions per crossing usually require re-weighting to match the collected data. Proper estimation of the number of pileup events is critical, as this has an effect on several kinematic parameters. For example, one of the most obvious is the measured amount of energy of a given jet. In an event with a high number of secondary pileup interactions, many low energy particles can end up in the primary interaction's jets, inflating the jet's true energy. For this example, jet energies typically have a correction that is applied which is dependent on the average number of interactions in that data-taking period [138].

6.1.2 Z Boson Mass Peak

The electronically decaying Z boson has a very clear signature in the ATLAS detector due to the production of two isolated leptons. As shown in Section 5.2.6, Z+jets events are produced at a high rate at the LHC. Additionally, when the Z decays to electrons, muons, or hadronically, the mass can be fully reconstructed, making for a good means of MC calibration and checks on normalization.

The Z boson control plots in Figure 6.1 show that the overall normalization agrees well between data and MC simulation for the Z+jets sample. The agreement shows that the pileup corrections which are used for all MC based samples are correct. Figure 6.1b suggests that at high jet multiplicity there is some disagreement. This is due to a known issue with the SHERPA generator [119] used to generate the Z+jets sample which requires some correction when there is a high jet multiplicity in an event. Furthermore, the number of

Table 6.1: Summary of all backgrounds in the Z boson enriched region.

Source	Count
Z	$1.189 \times 10^6 \pm 3.298 \times 10^3$
DiBoson	$4.174 \times 10^3 \pm 42.36$
$t\bar{t}$	$2.329 \times 10^3 \pm 23.18$
Single Top quark	$(4.17 \pm 0.08) \times 10^2$
$t\bar{t} + V$	19.4 ± 0.2

events drops by roughly a factor of eight for each additional jet, thus the total contribution of mis-modelled events is small. As this sample is not used in the signal region, no correction was applied. A summary of the backgrounds in the Z boson enriched region is given in Table 6.1.2.

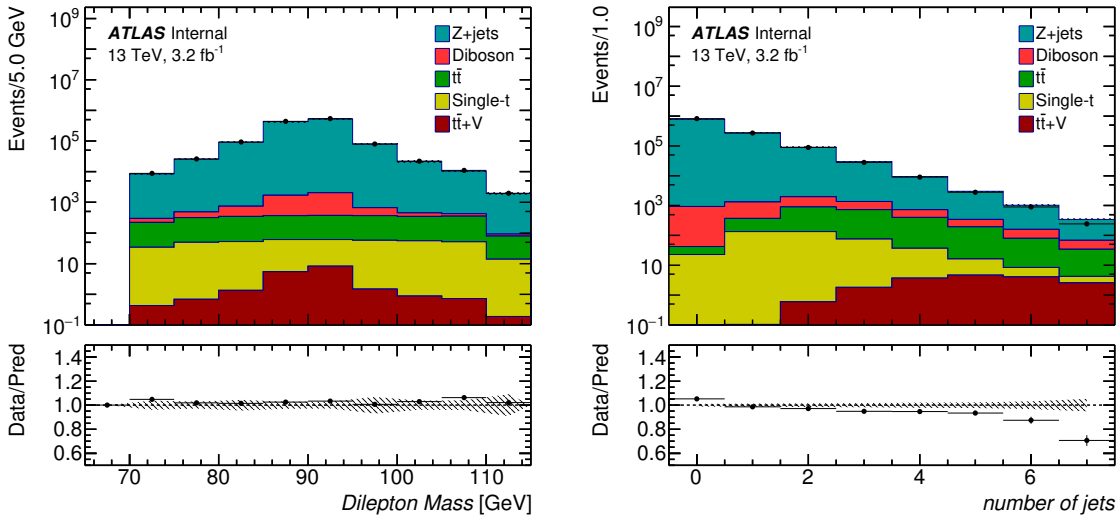
6.1.3 W Boson Mass Sidebands

As mentioned in Section 5.2, the only processes expected to contribute In the signal region are diboson and $t\bar{t} + V$. Regions enriched with diboson and $t\bar{t} + V$ are very difficult to specify without other more dominant backgrounds. This is due to the relatively low cross-section of the diboson and $t\bar{t} + V$.

In order to validate the data-driven backgrounds, it is necessary to examine the same sign region. However, to avoid examining the signal region inadvertently, creating a mutually exclusive region is crucial. In this case, it is simplest to examine the region where the invariant mass of the two jets is not close to the mass of the W boson. This region is called the W mass sidebands, and it is shown in Figure 6.2. This region is also sensitive to the diboson and $t\bar{t} + V$ backgrounds. This region again shows agreement between the data and backgrounds, even though it is quite statistically limited.

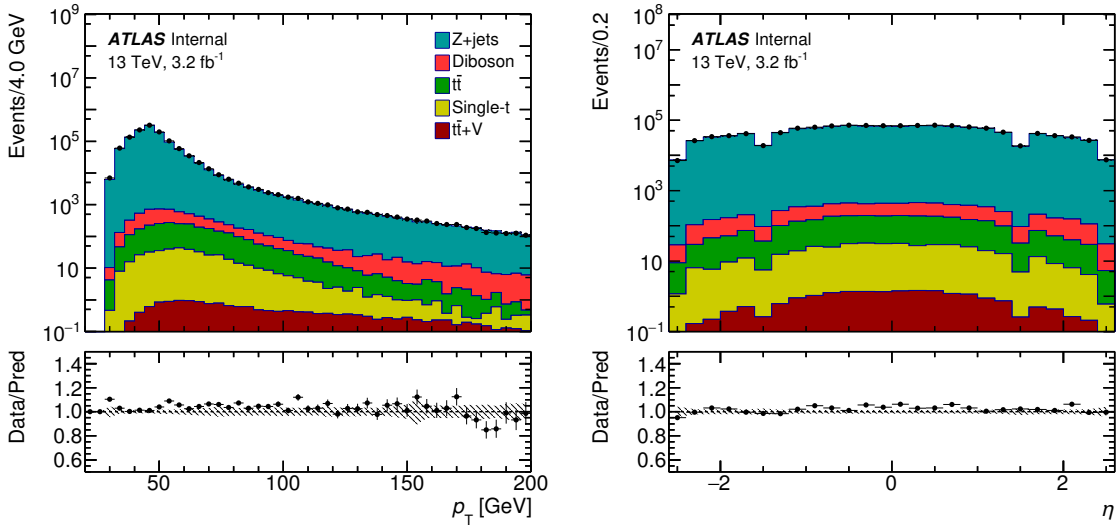
6.2 Signal-Background Separation

In order to specify the signal region a few kinematic cuts are made consistent with the decay seen in the Feynman diagram in Figure 5.1(a). Only two electrons are expected in



(a) Invariant DiLepton Mass

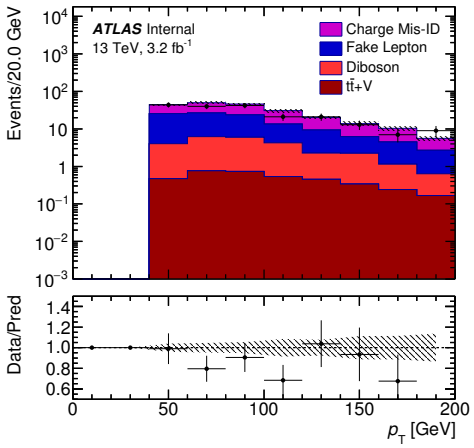
(b) Jet Multiplicity



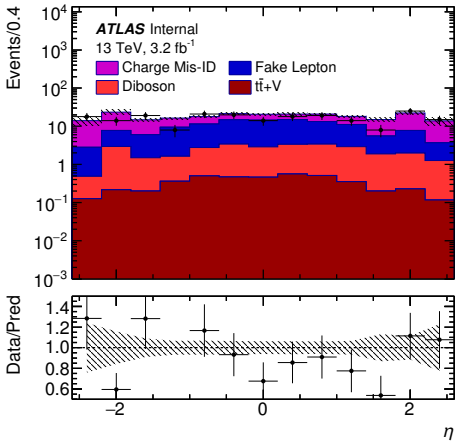
(c) Leading electron p_{T}

(d) Leading electron η

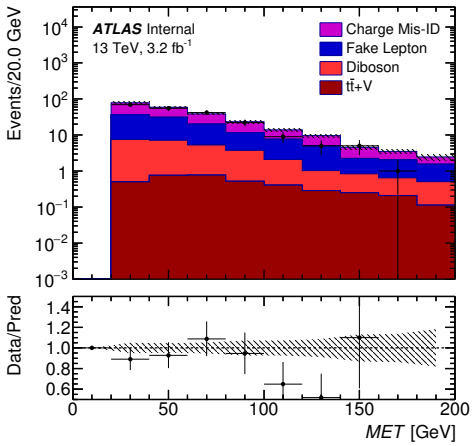
Figure 6.1: Opposite-sign Z control region, with DiElectron mass consistent with the Z boson. The hatched regions represent the statistical and systematic errors added in quadrature.



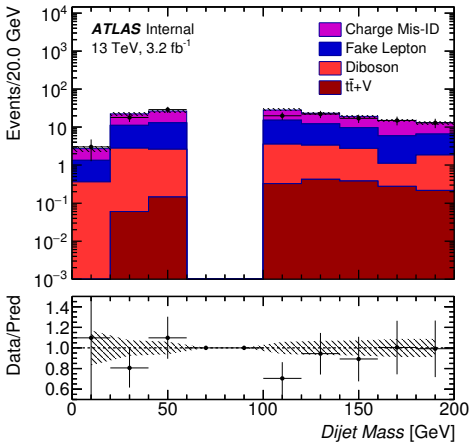
(a) Leading Electron p_{T} spectrum



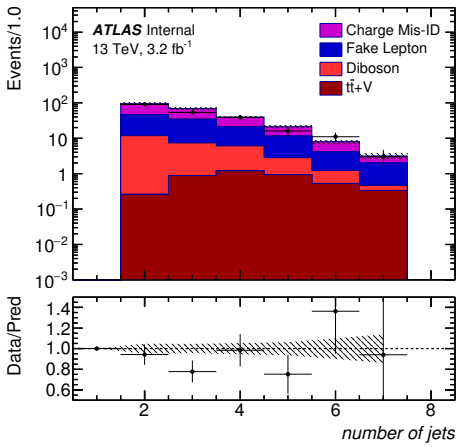
(b) Leading Electron η spectrum



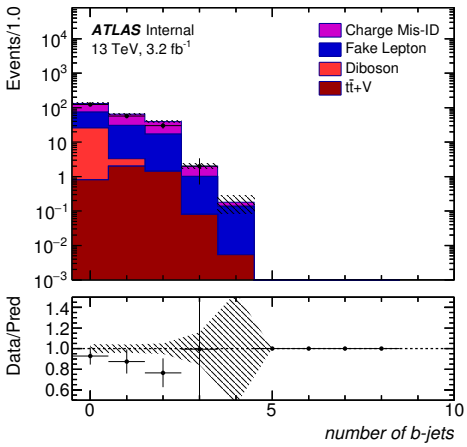
(c) $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum



(d) The Di-jet invariant mass spectrum



(e) Number of jets per event



(f) Number of b-tagged jets per event

Figure 6.2: Same-sign W boson mass sidebands control region. A cut excluding the dijet mass near the W boson mass was performed, $M < 60$ GeV and $M > 100$ GeV. The hatched regions represent the statistical and systematic errors added in quadrature.

the final state, so a restriction of only two electrons was imposed. Additionally, the W associated with the neutral heavy lepton is expected to decay hadronically to two quarks, which will form jets. By combining two jets in the event and requiring that the invariant mass of the parent particle is near the W boson mass, then the mass of the neutral heavy lepton can be fully reconstructed. This removes the majority of background events which have jets which do not originate from a W boson.

The electrons in the event obey the tight definition thereby reducing the contribution from fake electrons. An electron could be ill-defined because it passes through certain sections of the detector, such as the crack, or through dead modules. Events with these issues are discarded.

One of the most significant background rejection cuts is the same-sign electrons requirement. Physics processes with same-sign electrons are rare. In Figure 5.1(a), a given event has equal likelihood of having two leptons with the same charge as it does to have them with opposite charges.

The list of cuts is shown here:

- Veto events with ≥ 1 muons
- Select events with exactly 2 tight same-sign electrons
- Select events with ≥ 2 jets
- Invariant mass of the two highest p_{T} jets in the event to be within 20 GeV of the W mass
- $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV

Furthermore, in order to assure that the fake electron estimation is accurate in the signal region it was necessary to exclude the Z peak region. The fake rate is valid in regions where at least one of the electrons in the event is likely to be a fake electron. When the Di-lepton mass is near the Z peak, the electrons are likely to be real electrons which have been charge mis-identified. Therefore avoiding the Di-lepton peak is necessary to avoid double-counting

Table 6.2: Summary of the signal selection made in Figures 6.3, 6.4, 6.5, 6.6 for the 200 GeV Type III Seesaw heavy lepton.

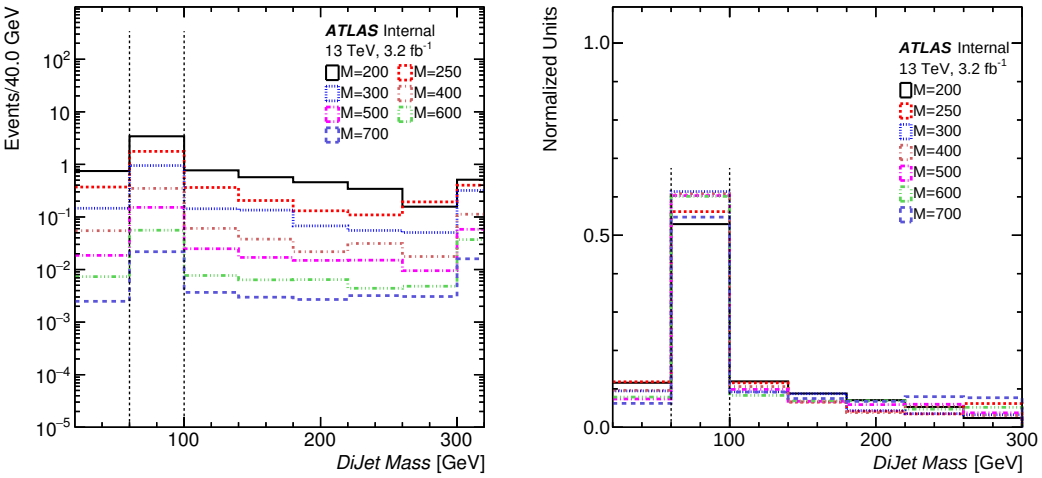
Selection	Number selected	Total Events	Percentage remaining
DiJet	3.44	6.98	49%
DiLepton	3.45	4.22	82%
Electron p_{T}	3.42	3.43	99.7%
$E_{\mathrm{T}}^{\mathrm{miss}}$	3.42	3.50	97.5%

contributions from the charge misidentification. An electron momentum cut is also applied to remove a region which has little signal, as seen in Figure 6.5:

- Di-lepton mass outside 20 GeV from the Z mass $|M(ee) - 91.2 \text{ GeV}| > 20$ GeV
- Leading electron $p_{\mathrm{T}} > 40$ GeV

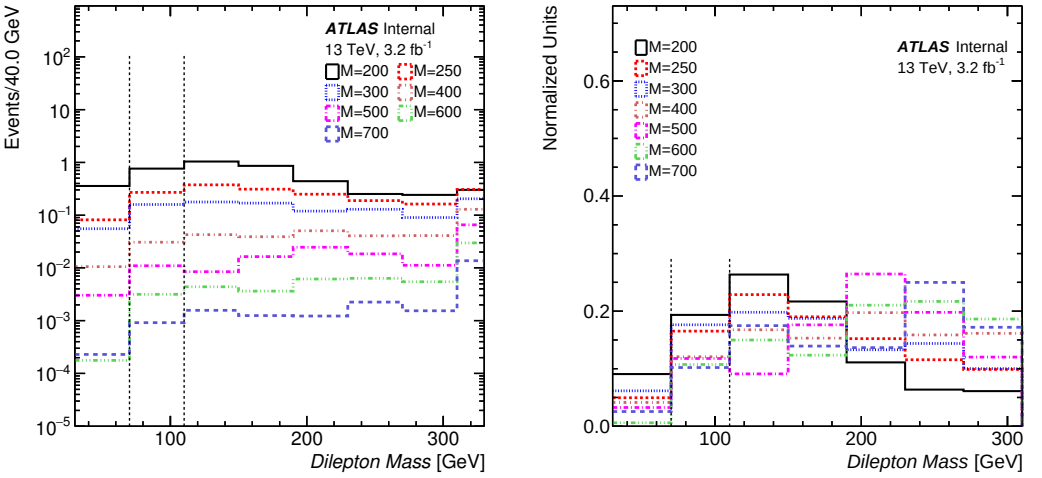
Figures 6.3, 6.4, 6.5, and 6.6 show four kinematic variable selections performed for the various mass point MC signal samples. Each figure shows the kinematic variable the selection is being made on with the other three cuts already applied. These are each plotted on a logarithmic scale to increase the visibility of smaller size samples. Additionally the right plots show the samples all normalized to equal area. The purpose of the normalization is to show the shape of the variables for each mass point.

In general these selections show that very little signal is lost due to the cuts. The most significant of the cuts is the DiJet mass cut. For the 200 GeV mass point, this cut removes 50% of the signal. However it is a crucial cut to ensure the jets originate from the W boson. The second most significant is the DiLepton mass selection, however as mentioned before, it is critical to ensure good modeling of the fake electron background. Table 6.2 shows a summary of the number of events and the total number which pass the independent selection requirements for the 200 GeV Type III Seesaw heavy lepton.



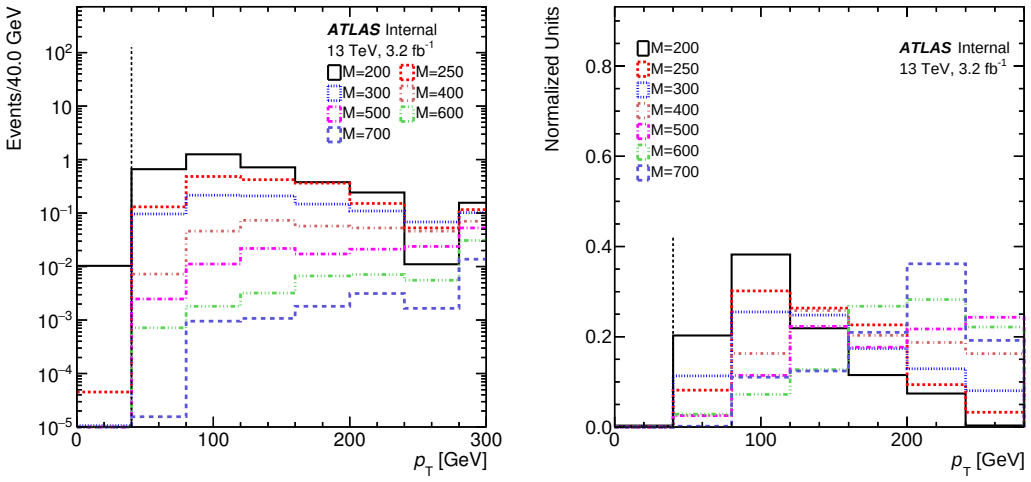
(a) Logarithmic scale (b) Samples normalized to equal area

Figure 6.3: Di-jet mass showing selection range around the W boson mass.



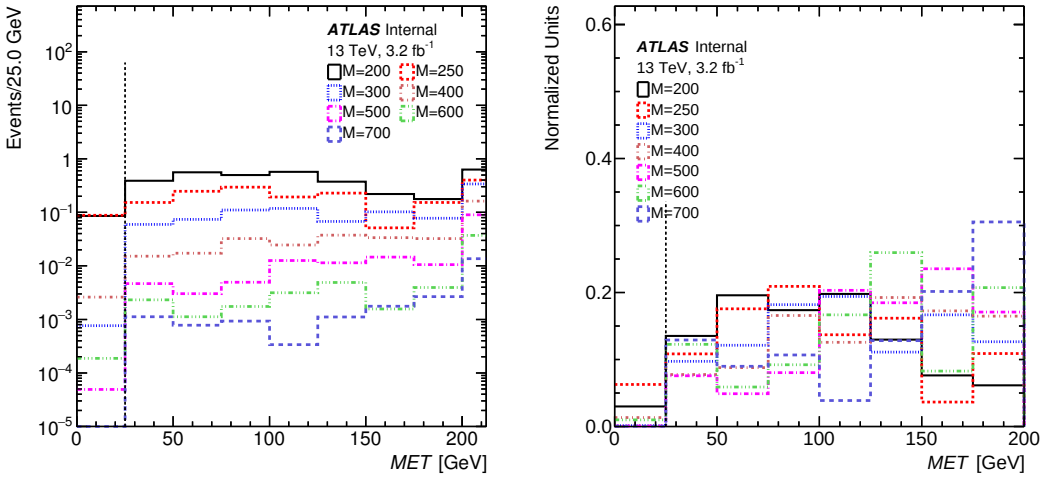
(a) Logarithmic scale (b) Samples normalized to equal area

Figure 6.4: Di-lepton mass distribution showing exclusion range around the Z boson peak. The Di-Lepton cut removes up to 18% of the signal events, however the cut is necessary to ensure the fake electron estimation is valid.



(a) Logarithmic scale (b) Samples normalized to equal area

Figure 6.5: Leading Electrons p_{T} distribution, showing the region < 40 GeV which will be removed. This cut removes only a few signal events but cuts away the low energy background.



(a) Logarithmic scale (b) Samples normalized to equal area

Figure 6.6: $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of selected simulated signal events. All selection requirements have been applied except for $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV. The $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement removes at most 2.5% of the signal events depending on the heavy lepton mass examined.

6.3 Background Summary

At this stage of the analysis a fairly stringent set of cuts have already been specified. The greatest reduction in events is due to the use of the same-sign selection. After these selection requirements have been applied the background events which remain are primarily due to detector effects, namely fake electrons and charge mis-identified electrons. Table 6.3 shows the total count of the background sources for the signal region before further selections.

Table 6.3: Summary of all backgrounds with statistical and systematic errors after event selection.

Source	Count	Statistical Error	Systematic Error
DiBoson	6.9	± 1.4	$^{+0.9}_{-1.1}$
$t\bar{t} + V$	0.8	± 0.1	$^{+0}_{-0.1}$
Fake Electrons	22.5	± 1.8	$^{+3.3}_{-2.2}$
Charge Mis-ID	28.2	± 1.6	$^{+8.1}_{-7.0}$
Total	58.4	± 4.9	$^{+12.3}_{-10.35}$

6.4 Cut-Based Optimization

In the Run 1 analysis, the technique used to separate the signal from the background was an optimized set of kinematic cuts. The variables utilized were:

- $E_{\mathrm{T}}^{\mathrm{miss}}$
- Leading Electron p_{T}
- Second Leading Electron p_{T}
- Leading Jet p_{T}
- Minimum Di-Lepton mass
- dR between the two jets

These variables were chosen based on their ability to discriminate between the signal and backgrounds. In order to determine at what value to cut on these variables, an optimization algorithm was run. The algorithm looped through roughly 10 different cut values

for each of the variables to be tested. The surviving expected number of events of signal and background, were then used to calculate a sensitivity. Typically a figure of merit chosen for this kind of optimization is $\frac{S}{\sqrt{B}}$, where S is signal and B is background. However in the case were $S \approx B$, it is more appropriate to use Equation 6.1 [139].

$$\sqrt{2 \times (S+B) \times \log\left(1+\frac{S}{B}\right) - S} \tag{6.1}$$

In order to make a comparison of techniques, the same optimization method was performed in this analysis. The optimization was run separately for each of the generated signal mass points [GeV]: 200, 250, 300, 400, 500, 600, 700. The results are sets of cuts for each specific signal mass point, listed in Table 6.4. After the cuts the total number of signal events and background events were used to measure the sensitivity of this analytical method.

Table 6.4: Summary of optimized kinematic cuts for the various signal mass points.

$m(\Sigma)$	min $E_{\mathrm{T}}^{\mathrm{miss}}$	min DiLepton Mass	min $p_{\mathrm{T}}1$	min $p_{\mathrm{T}}2$	min Jet p_{T}	dR Jets
GeV/c^2	GeV	GeV	GeV	GeV	GeV	
200	100	110	60	25	25	2.5
250	130	170	50	50	50	2.5
300	140	150	70	50	40	2.5
400	140	190	70	25	50	2.5
500	140	190	70	25	90	1.5
600	140	190	70	40	80	2.0
700	140	190	50	50	120	2.5

Ultimately, the performance difference between the cut combinations for each mass point seen in Table 6.4 is minimal. The significance of each optimized cut is shown in Table 6.5. σ is the significance calculated from the number of signal events over the square root of background events. This figure of merit is used rather than Equation 6.1, as the number of signal and background events is not similar when using optimized cuts.

The majority of the discriminating power comes from the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut which is explained in more detail in Section 7.3.2. In Run 1, only one cut combination was chosen for all mass

Table 6.5: Summary of significance for optimized heavy lepton selection.

Mass point optimized	σ_{200}	σ_{250}	σ_{300}	σ_{400}	σ_{500}	σ_{600}	σ_{700}
200 GeV	1.04	0.55	0.38	0.18	0.08	0.03	0.01
250 GeV	0.91	0.85	0.69	0.37	0.18	0.08	0.03
300 GeV	0.88	0.76	0.71	0.35	0.18	0.07	0.03
400 GeV	0.86	0.74	0.69	0.40	0.21	0.09	0.04
500 GeV	1.61	0.85	1.21	0.62	0.42	0.19	0.07
600 GeV	1.41	0.97	1.31	0.75	0.43	0.19	0.08
700 GeV	0.75	0.86	0.99	0.52	0.32	0.16	0.08

points to simplify the analysis and because there was not a great difference between the results when the combinations were compared. To compare to the previous analysis the same was done here. The 300 GeV mass point cut combination was chosen as it is closest to the mass limit set in Run 1.

6.5 Multivariate Techniques

Another, more sophisticated method of separating the background events from signal is to use machine learning to optimize the selection. Machine learning requires training, which is optimizing a machine learning algorithm by inputing data representative of the system to be modeled. Two major problems which machine learning is used to solve are regression, and classification. This analysis focuses on using classification to identify signal and background events. Machine learning is effective at identifying differences that may be non-obvious to human analyzers by examining multi-variate correlations. The methods are not limited to cuts on variables, but possibly can use complex combinations of variables to achieve higher discriminating potential. Two types of multivariate techniques were explored in this analysis: boosted decision trees and neural networks.

6.5.1 Boosted Decision Trees

A decision tree is a method of to classify data into different groupings. For the purpose of this analysis, it is only necessary to separate the information into two types: signal and background. An algorithm is used to go through all the available variables and find the optimal splits to make for each of the variables. This split is called a node. There are several notable algorithms (ID3, C4.5, CART) [140, 141] for making decision tree splits based on input variables.

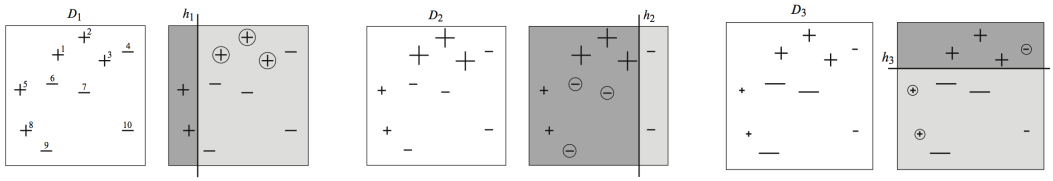
The criterion for determining the best split is done using the signal purity in the resultant children nodes. For CART, the algorithm tries to minimize Gini impurity, defined in Equation 6.2, for both children.

$$\text{Gini} = \left(\sum_{i=1}^{n} W_i\right) P(1-P) \tag{6.2}$$

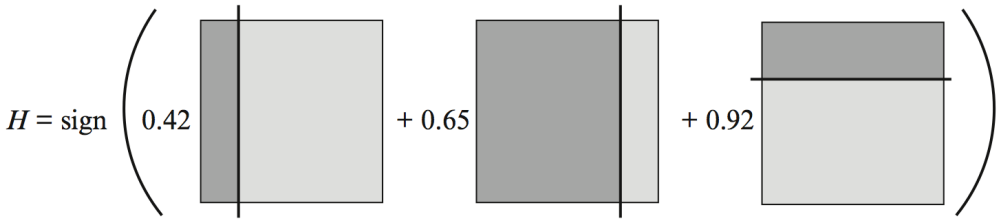
Where W_i is the weight of the data point, P is the probability of correctly classifying the point, and $1-P$ is the probability of mis-classifying the point. G_{ini} reaches 0 when all points fall into a single classification.

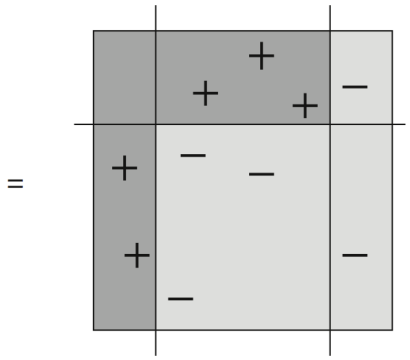
This is done over many nodes, and the final resultant children are called leaves. Any leaf which has a signal weight of more than 50% of the total weight is classified as a signal leaf; otherwise it is classified as a background leaf. A decision tree is a form of cut optimization, and by itself is not stable to changes or fluctuations in the data.

Simple decision trees are not good models of complex data and will have misclassified events. For example, signal events can be classified in background leaves and vice versa. Boosting increases the ability to discriminate signal from background over a simple decision tree. Boosting generates thousands of different decision trees and adds them together. Boosted decision trees (BDTs) are trained using a set of data, which optimizes the performance of the BDT. Subsequently an orthogonal set of data can be examined with the trained BDT to categorize the data into signal and background categories. There are several methods of boosting, two of which are Adaboost, and gradient boost.



(a) Tree 1 with equal weighting (b) Increased weights on misclassified events in Tree 1 (c) Increased weights on misclassified events in Tree 1,2





(d) By combining all the decision trees with weights depending on the performance of the discriminant, a much better classifier can be created.

Figure 6.7: The Adaboost method to boost decision trees [142]. h_i are classifiers, D_i are distributions of data, and H is the weighted combination of classifiers h_i. The weights are the result of an algorithm based on the error e_i of each h_i.

6.5.1.1 AdaBoost [142]

There are several means of determining a weight value, one of the most commonly used methods is called AdaBoost. Figure 6.7 shows the method applied to a simplified two dimensional case. Trees are developed using Gini on the input data, typically very short decision trees are used, possibly with only a single split (stumps). The trees are commonly constrained by specifying a maximum number of splits, leaves, or layers. The children signal and background leaves are evaluated for purity and mis-classified events are re-weighted to have more significance. The re-weighted input data is used to generate a new decision tree and classification is re-evaluated. The cycle is repeated a few hundred times, and one ends up with a set of decision trees, also known as a forest. The forest is combined together with weights based on the tree's effectiveness to give a final set of classification weights.

6.5.1.2 Gradient Boost [143]

Another method commonly used is called gradient boost. Rather than re-weighting data points, this method works using decision trees with the help of the minimization of a loss function. With the Adaboost method, misclassified events are re-weighted, while with the gradient method a loss function is evaluated on the leaves calculating how much the decision tree model deviates from the input data.

An analogous description of the gradient boost concept follows. Let $F(x)$ be a function to classify input information, and y is the information being classified, then $y - F(x)$ is the residual. In gradient boosting, a new hypothesis $h(x)$ is trained on the residual $y - F(x)$, such that the functions $F(x) + h(x)$ are more consistent with y, and a better model for the relevant data. Further hypotheses $h_i(x)$ can be fit to the residuals of $F(x) + h(x)$. Fitting $h(x)$ to a residual is equivalent to using a negative gradient of a loss or error function.

The gradient boosting requires defining a differentiable loss function. Many decision stumps are created from the input data. A single tree is used as $F(x)$ and further trees which minimize the loss function are added to $F(x)$ iteratively. Existing trees in the model

$F(x)$ are not changed. Once a fixed number of trees are added or no significant improvement of the model is seen, training stops. For classification, signal and background errors are calculated independently and regression trees are fit to the negative gradient of the errors of each categorization.

6.5.2 Neural Networks

Neural Networks are another means utilized to better understand the differences between a signal and a background. The method is modeled after human thought. Essentially taking multiple variables into account and making a decision based on a combination of those variables. Early in the history of neural networks, a type of artificial neuron was used called the perceptron. It takes as input several binary variables and produces a single binary output. If the weighted sum of the inputs is greater than some threshold, then the perceptron outputs true, otherwise it outputs false. This is analogous to weighing evidence to make a decision. In contemporary use, the perceptron is replaced with a different type of neuron that takes as input and outputs a non-integer value between 0 and 1.

6.5.3 Machine Learning Complications

There are several problems which are related to improper use of machine learning:

- Over-fitting
- Outliers
- Noise

Over-fitting is the result of training the algorithm too specifically on the training sample. To demonstrate this issue examine Figure 6.8. The three graphs show the same dataset depicted with blue dots, where the true data has a parabolic dependence. The data points are smeared around the true parabolic values by Gaussian noise. The black lines in the figures are attempts to fit the data points. On the left the line is under-fit, it does not fit

the shape of the data. The central figure is a good fit. The right figure is the result of over-fitting. It represents this particular data very well, but additional data will not follow this pattern and will result in poor performance overall.

In order to evaluate if the training is successful, the training and testing samples are compared. The training and testing samples should be constructed by splitting the available data into orthogonal samples. The output of the machine learning method used in this analysis is a discriminant value for each event based on several kinematic variables for that particular event. By comparing the histograms of the discriminants of the training samples versus the testing samples, and looking for good agreement, it can be determined that there is no over-fitting. The testing events would tend to follow the parabola, and the over-fit curve would very poorly represent them. Thus, while the discriminant for the training sample might be sharply peaked at 'very signal-like', the testing data is not likely to share that behavior.

Additionally, one could imagine a case where a data point could exist far from the parabola. This would be an example of an outlier; a data point which does not follow the overall trend of the data. It is not desirable that outliers influence the training of the algorithm, as outliers are not consistent with the patterns of the data.

Noise is a familiar concept in all experimental work. If the Gaussian noise in the leftmost plot of Figure 6.8 is large, the pattern of the data will not be discernible.

6.5.4 Machine Learning in the Analysis

The background and signal MC samples are each separated into two equal sized subsets: one for use in training and the other for testing. The sample is split randomly to avoid potential correlations in the ordering of the datasets. The training samples are not used afterwards in the analysis, as events that are used to train the algorithm have a bias. The testing samples are kept and their total weights were increased by a factor of two to make up for the lost events in the training sample.

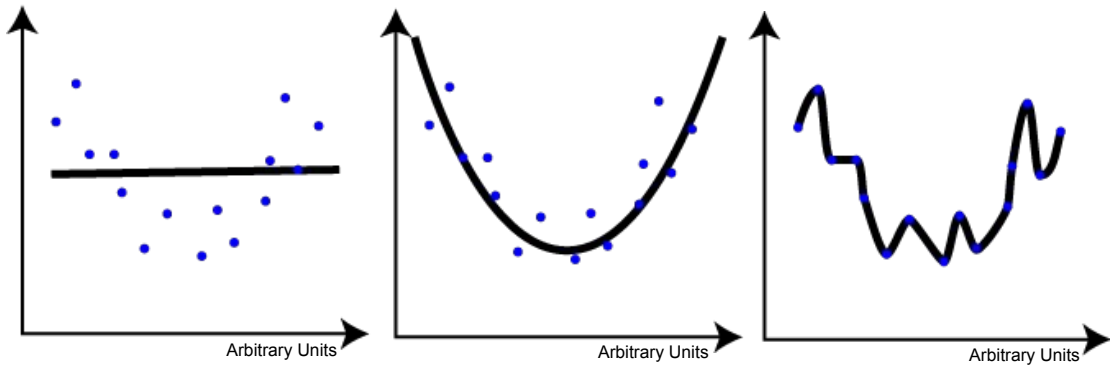
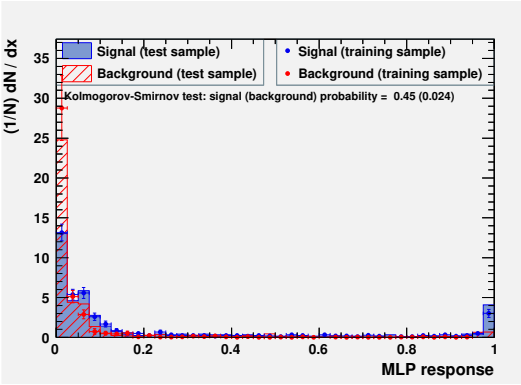


Figure 6.8: The blue points have a parabolic dependence. The black lines are attempts to represent the pattern of the data. The left most figure shows under-fitting, the middle line shows a good fit, and the last line is over-fit [144].

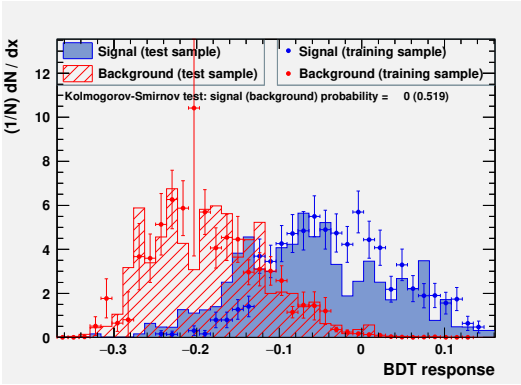
To check if a method is over-fit, it is recommended to check the correlation of the discriminant between the training and testing sample. If they are in good agreement, it is unlikely the method is over-fit. This can be done visually or via a Kolmogorov-Smirnov test. The three methods of machine learning previously described have been tested for this analysis: Adaboost, gradient boosted decision tree (BDTG), and a neural network. For these methods Figure 6.9 shows the comparison of testing and training for the output discriminants. The figures also show p-values for the Kolmogorov-Smirnov tests, the Adaboost method fails the test with a p-value of 0.

A toolkit called TMVA was used for the implementation of the machine learning methods [145]. Ultimately a gradient boosted decision tree (BDTG) is used, which produces a discriminant with a smooth distribution, with a consistent shape between the training and testing samples. The variables used to train the BDTG discriminant are:

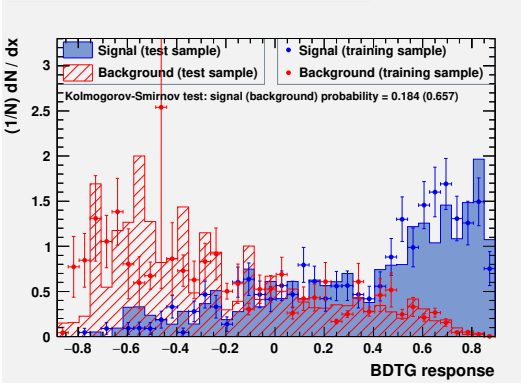
These variables are chosen for their good agreement between background estimation and data in the control regions. The variables used for the cut optimization are also included here as the signal MC shows they have distributions which are distinct from the dominant backgrounds. The electron plus di-jet variables are included for the same reasons, as they



(a) Neural Network



(b) Adaboost BDT



(c) Gradient BDT

Figure 6.9: Normalized distributions of the multivariate output for signal and background. The agreement between the training and testing distributions indicates that the discriminants are not over-fit.

Table 6.6: Summary of the variables used to train the BDTG discriminant and their names in subsequent figures.

Variable	Machine learning variable
Leading electron p_{T}	tmva_LeadingLeppt
Second leading electron p_{T}	tmva_secondLeadingLeppt
Di-lepton mass	tmva_dilepton
Leading jet p_{T}	tmva_LeadingJpt
Second leading jet p_{T}	tmva_secondLeadingJpt
Di-jet mass	tmva_dijet
dR between the two leading jets	tmva_dRjet12
Leading electron + di-jet mass	tmva_heavy
Second leading electron + di-jet mass	tmva_notheavy
dR between the leading electron and the di-jet object	tmva_dRheavy
dR between the second leading electron and the di-jet object	tmva_dR_not_heavy
$E_{\mathrm{T}}^{\mathrm{miss}}$	tmva_Met_et

have sensitivity to the signal model. Many of the variables do not discriminate between the signal and background at low mass, however they do at larger heavy lepton masses.

The BDTG discriminant is broadly distributed between -1 and $+1$, with most events near $+1$. At high mass points the discriminant only exists in a sharp peak at 1, and the background is cleanly separated. This suggests that the multivariate technique can more easily distinguish between the samples, and in fact, may not be necessary.

Figure 6.11 shows the input variables used for the BDTG for the 700 GeV mass point. The background shapes are identical to Figure 6.10. The energy available in the signal events allows much better separation between signal and background. The electron and jet p_{T} variables show a more significant shift to higher momenta, and the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is shifted to higher values. The heavy lepton mass distribution shows a very distinct separation. With more data, it may be advantageous to make a simple cut on heavy lepton mass, however this will need further study.

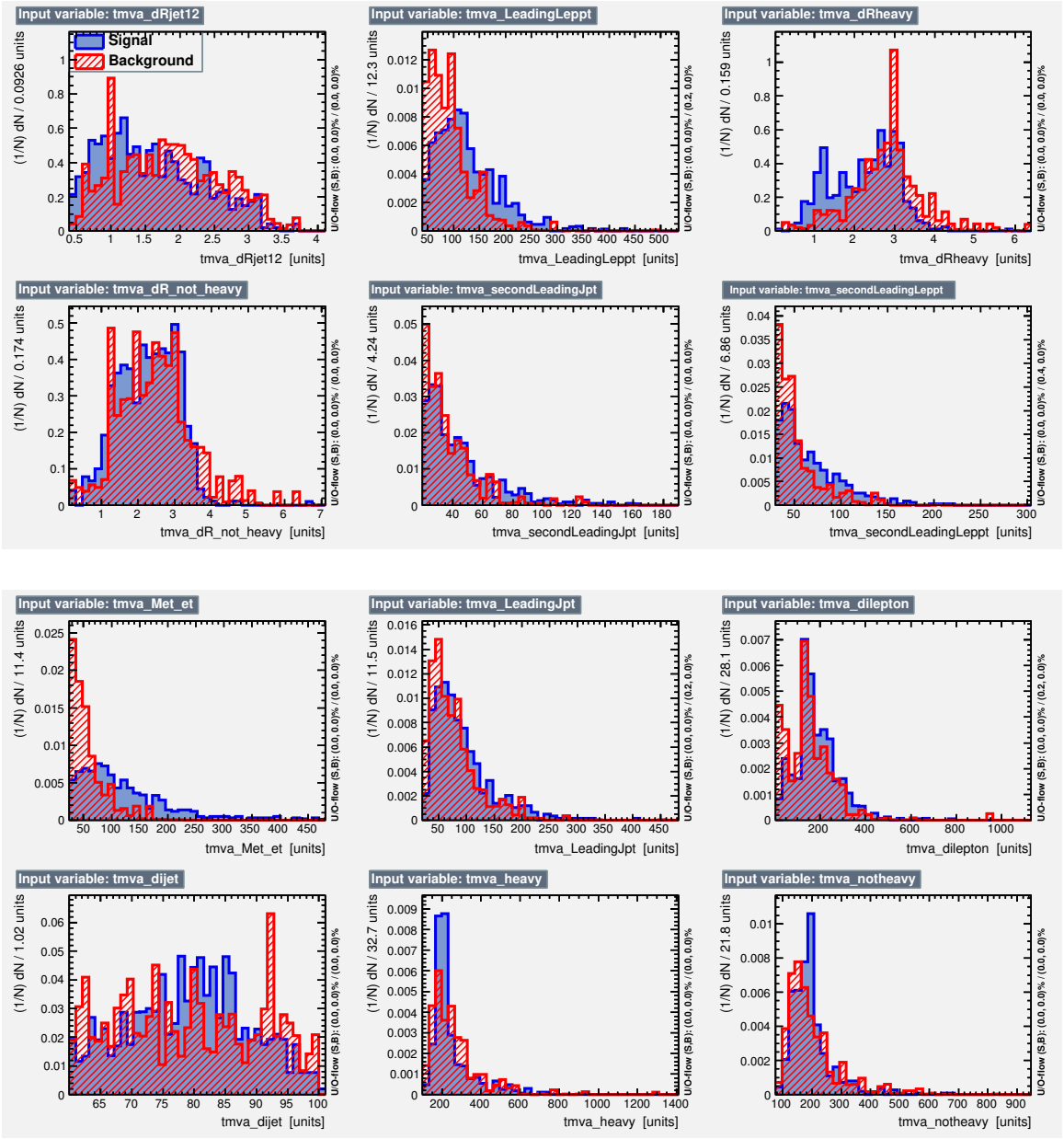


Figure 6.10: Input variables used in the machine learning for the 200 GeV mass point. See Table 6.6 to translate the variable names.

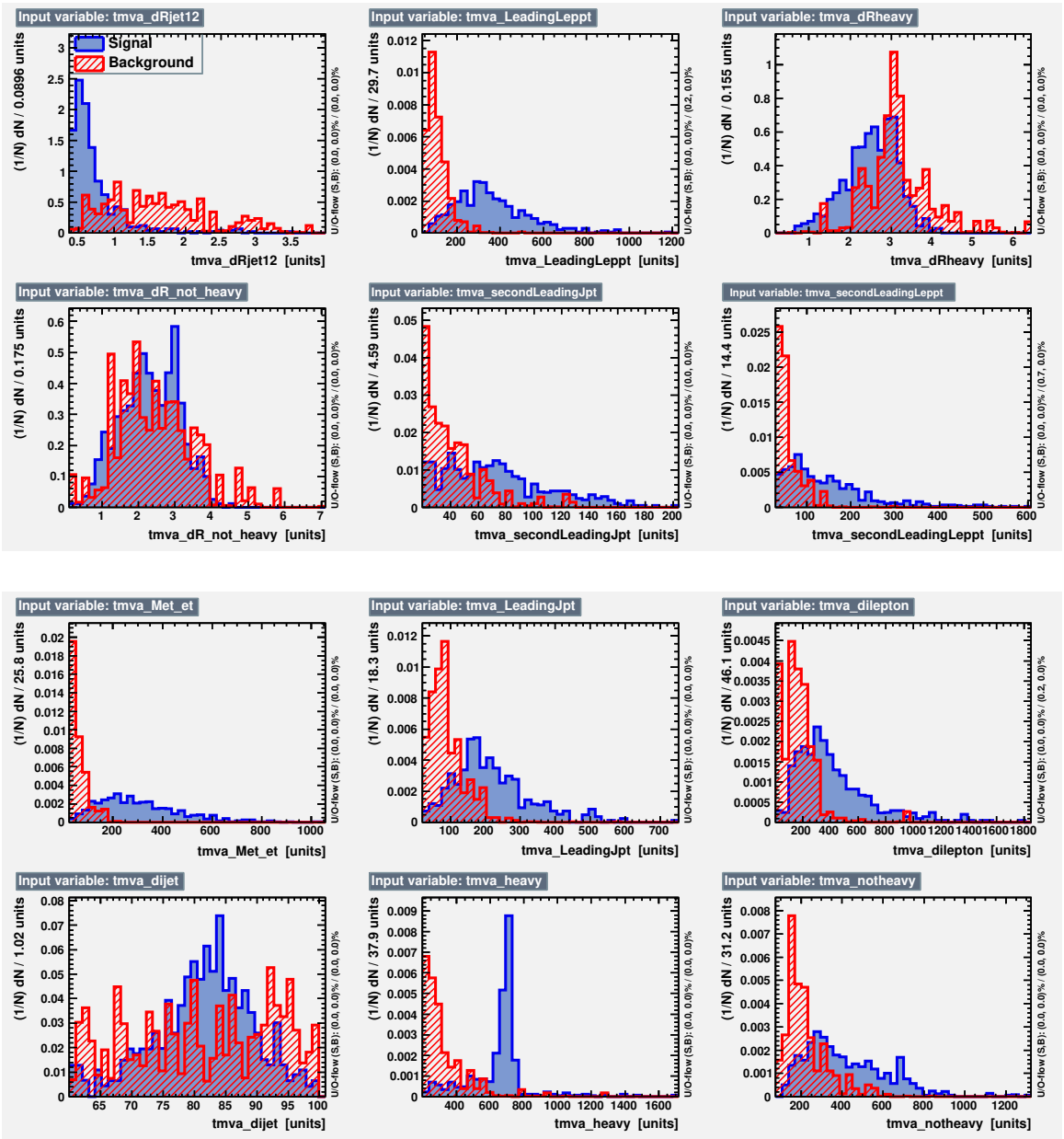


Figure 6.11: Input variables used in the machine learning for the 700 GeV mass point.

CHAPTER 7. RESULTS

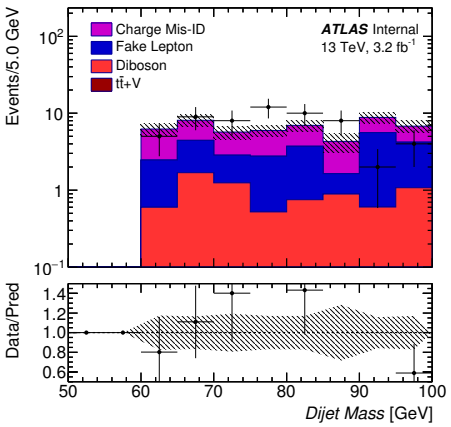
This chapter presents the results of the analysis. The two different methods described in the previous chapter are used to select events in the signal and background control region. These events are evaluated using frequentist methods to determine for which heavy lepton mass the signal model can be excluded.

7.1 Boosted Decision Tree Discriminant Validation

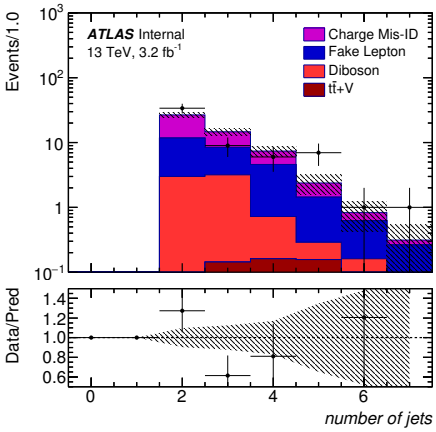
In order to make sure that the background is correctly described in the signal region before unblinding, a background region is examined. The BDTG discriminant distribution is split into two regions, a background-like BDTG region and a signal-like BDTG region. The background-like BDTG region is defined as the region between -1 and the split value, and the signal-like BDTG region is from the split value to $+1$. The choice of the split value depends on the signal mass point. The signal-like BDTG region is the signal region where the cross-section measurement is made.

Figure 7.1 shows a few kinematic plots comparing data and background in the background-like BDTG region.This region is expected to have very little to no signal and thus suited for examining agreement between data and background. These particular plot are shown for a BDTG trained with the 200 GeV heavy lepton mass signal. The data and simulation show an overall good agreement in the background-like BDTG region.

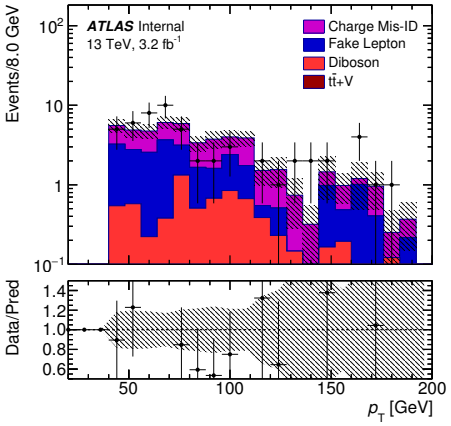
After confirming the data and background estimation has good agreement in the background-like BDTG control region, the signal region is examined. There is no obvious signal observed in the signal region. In the absence of a signal, we calculate a limit on the heavy neutrino pair production cross-section.



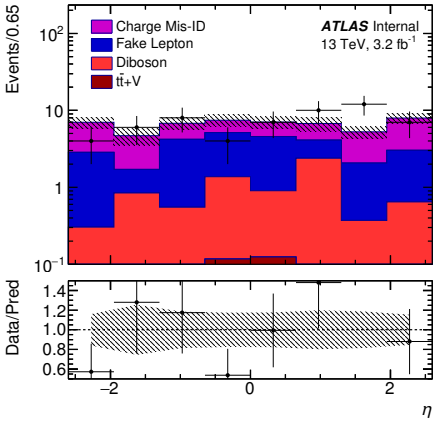
(a) Invariant DiJet Mass



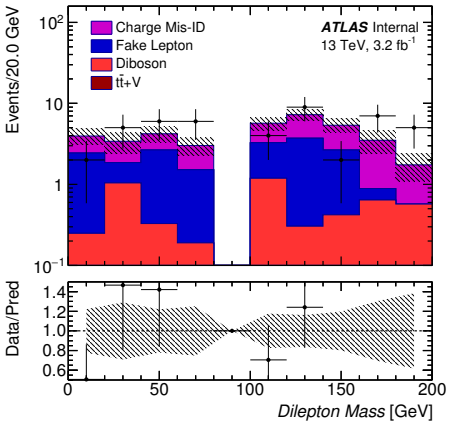
(b) Jet Multiplicity



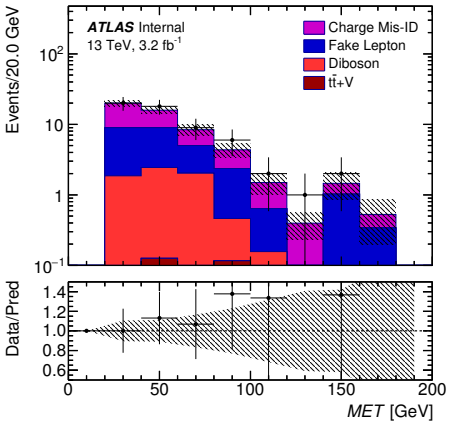
(c) Leading electron p_{T}



(d) Leading electron η



(e) Di-Electron Mass



(f) $E_{\mathrm{T}}^{\mathrm{miss}}$

Figure 7.1: BDTG background-like control region for the 200 GeV heavy neutrino mass selection showing good agreement between data and background simulation. This region has the same pre-selection requirements as the signal region, except that it requires BDTG < 0 (background-like region).

7.2 Systematic Uncertainties

Several algorithms are used in this analysis which allow for calibration of the various objects such as electrons, jets, and muons. These tools are continuously curated and maintained by the ATLAS collaboration based on the output of the detector to ensure that the particles observed are being correctly reconstructed. Additionally these algorithms perform the cleaning of jets, the reconstruction of $E_{\mathrm{T}}^{\mathrm{miss}}$ in a given event and the calculation of the likelihood that a given jet originated from a b quark.

All of these algorithms also provide the uncertainties on their corrections, which must be taken into account in the analysis. These errors represent the uncertainty due to imperfect modeling and approximations. Statistical uncertainties can always be improved by analysing larger amounts of data, however systematics uncertainties are inherent to the method used.

Several systematic uncertainties are only pertinent to Monte-Carlo simulated events, and therefore do not contribute greatly in this analysis. The major uncertainties in this analysis have already been discussed in Sections 5.4.1 and 5.4.2. The uncertainties on the fake factor method and the method used to calculate the rate of electron charge Mis-ID are large compared to the errors on the Monte Carlo backgrounds.

For all the histograms shown in this document the systematic uncertainties were added in Gaussian quadrature [146], in order to compute the total systematic uncertainty. This was then added in quadrature to the statistical uncertainty to give the total error shown in the figures.

7.2.1 Data-driven Background Estimation Uncertainties

The most significant uncertainties are due to the methods employed to calculate the fake and charge Mis-ID contributions. A summary of the data-driven uncertainties for the 200 GeV heavy neutrino selection is shown in Table 7.1.

Table 7.1: Uncertainties due to data-driven estimations. There are 12.8 expected events in the BDTG signal-like region for the 200 GeV heavy neutrino selection.

Systematic	Percentage error
Fake electron method	6.03 %
Charge Mis-ID method	11.93 %

7.2.2 Luminosity

The total integrated luminosity delivered to the ATLAS detector from the LHC has some associated uncertainties. If the luminosity is mis-measured it would manifest as an overall normalization error. This systematic results in a 2.1% uncertainty in the luminosity measurement [147].

7.2.3 Electron Selection and Reconstruction

Uncertainties regarding electrons are relevant for the Monte Carlo background estimates used in this analysis. There are uncertainties with the electron energy scale and resolution [148]. There are uncertainties associated with the single electron trigger efficiency [149]. There are also uncertainties with the reconstruction and with the measurements of the isolation parameters [150]. Variations are provided by the tools used to correct the corresponding parameters, and the variations are used in the total error calculations. A summary of the electron related uncertainties for the 200 GeV heavy neutrino selection is shown in Table 7.2.

7.2.4 Muon Selection and Reconstruction

Muon related uncertainties only show up in this analysis due to the applied muon veto. In particular, at low p_{T}, the muons can be mis-measured. In the case of this analysis, muons can be identified as a veto muon or discarded depending on this p_{T}. By running variations using the muon tools provided by physics analysis groups, these cases can be tested and

Table 7.2: Uncertainties due to electron measurements. There are 12.8 expected events in the BDTG signal-like region for the 200 GeV heavy neutrino selection.

Systematic	Percentage error
Electron energy resolution	0.1 %
Electron energy scale	0.71 %
Electron ID efficiency	0.38 %
Electron isolation efficiency	0.29 %
Electron reconstruction efficiency	0.15 %
Electron trigger efficiency	0.01 %

used as uncertainties. These uncertainties were not significant, and are not listed in the summary tables.

7.2.5 $E_{\mathrm{T}}^{\mathrm{miss}}$ Estimation

$E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainties in this analysis primarily originate from soft track measurements which are used for reconstructing the $E_{\mathrm{T}}^{\mathrm{miss}}$. The dominant uncertainties are energy scaling and resolution [151]. A summary of $E_{\mathrm{T}}^{\mathrm{miss}}$ related uncertainties for the 200 GeV heavy neutrino selection is shown in Table 7.3.

Table 7.3: Uncertainties due to $E_{\mathrm{T}}^{\mathrm{miss}}$ measurements. There are 12.8 expected events in the BDTG signal-like region for the 200 GeV heavy neutrino selection.

Systematic	Percentage error
$E_{\mathrm{T}}^{\mathrm{miss}}$ soft track resolution smearing parallel	0.02 %
$E_{\mathrm{T}}^{\mathrm{miss}}$ soft track resolution smearing perpendicular	0.05 %
$E_{\mathrm{T}}^{\mathrm{miss}}$ soft track scale	0.12 %

7.2.6 Jet Selection and Reconstruction

Jet uncertainties at ATLAS are primarily related to energy resolution and scale [106,152]. These uncertainties are grouped into sets concerning different nuisance parameters. There are also errors associated with JVT [104], however they were smaller than 0.005% in this

analysis. A summary of jet related uncertainties for the 200 GeV heavy neutrino selection is shown in Table 7.4.

Table 7.4: Uncertainties due to jet measurements. There are 12.8 expected events in the BDTG signal-like region for the 200 GeV heavy neutrino selection.

Systematic	Percentage error
Jet high p_{T} nuisance parameter	1.27 %
Jet nuisance parameter from in-situ analyses	1.37 %
Jet eta intercalibration	0.05 %

7.3 Cross-section Limits

Limit setting was performed with HistFitter [153] which utilizes Roostats [154]. The Roostats package provides many high-level statistical tools for confidence interval estimation and hypothesis testing. It was used in this analysis to test the background-only hypothesis to provide a limit on the Type III Seesaw Heavy Lepton mass.

In the cut-based method, a simple one bin measurement was used to calculate the expected limit of the analysis method. Any event which passed the cuts for a particular mass point was counted and the number of signal and backgrounds events was compared. The data yield was assumed to be equal to the expected number of background events. This is called an exclusion fit.

In order to calculate the significance of the exclusion, a frequentist approach was utilized. The frequentist school of thought is equivalent to calculating the probability of the data given the hypothesis. It essentially calculates the frequency of various outcomes. Specifically for this analysis, 10,000 toy models were used to calculate the confidence intervals for each tested value of μ_{Sig}. μ_{Sig} is a weight factor applied to the signal yield. The 95% confidence of excluding the signal is calculated by testing if the signal yield times μ_{Sig} plus background is consistent with the observed data. The largest value of μ_{Sig} for which the prediction is consistent with observation, at 95% confidence, is the observed limit of the model.

For the multivariate method a more sophisticated approach was applied. Rather than a single bin, several bins of the BDTG discriminant were used to perform the same measurement as described above. The number of bins and their size were calculated by analyzing the amount of statistics available and maximizing the signal to background ratio. This is an optimization technique used to gain a larger sensitivity than is possible with a single bin or multiple same-sized bins. Again the simulated events are fit to the level of the expected background, and the expected sensitivity of the measurement was calculated. As the BDTG background-like control region has a consistent background estimation and the binning is decided, the signal region is unblinded.

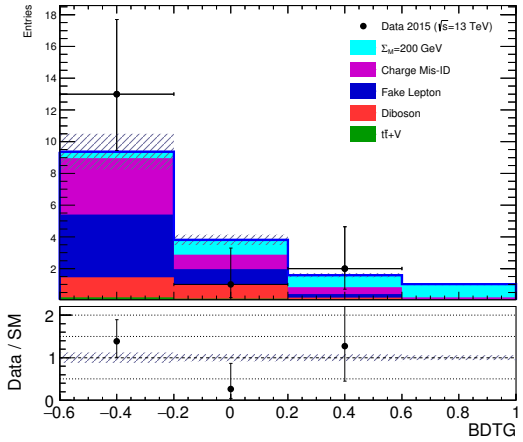
Figure 7.2 shows the bins used for each mass point and the total contributions from observed data, and predicted background and signal. Figures 7.2b and 7.2c show that there is a slight excess of observed data in the signal region. This is reflected in the limit setting in Figure 7.3a discussed in the next section. It is likely that this excess is a fluctuation. In total it amounts to less than two standard deviations of significance of the data above expectation. On average, this should happen about 1 in every 20 experiments. All other mass points are quite consistent between the data and the background-only hypothesis.

7.3.1 Results

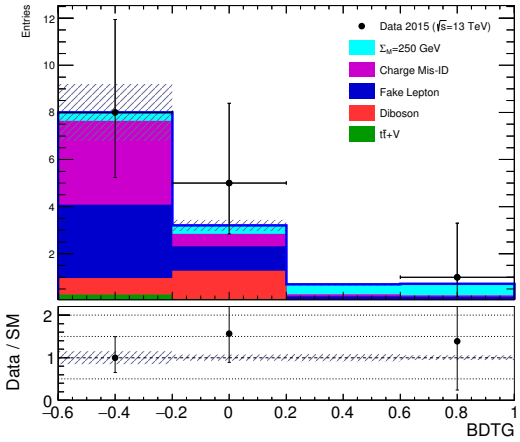
By using HistFitter, values for μ_{Sig} of the signal were calculated. This was done for the cut-and-count method and for the BDTG method. The μ_{Sig} can then be scaled by the cross-section and the k-factors and generator filter efficiency, introduced in Section 5.2, in order to calculate the exclusion cross-sections. The exclusion limit versus heavy lepton mass is shown in Figure 7.3a.

$$\sigma_{\text{Excl}} = \sigma_{\text{Gen}} \times k \times f_{\text{Gen}} \times \mu_{Sig}$$

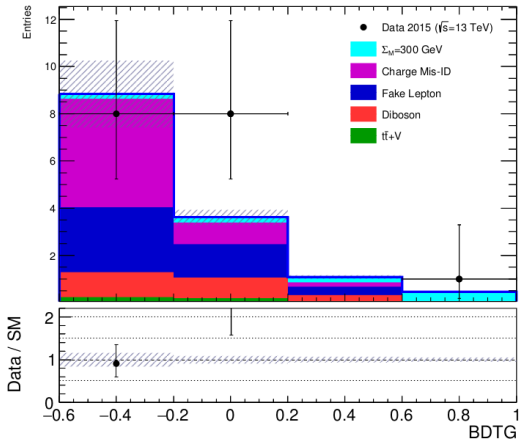
Additionally, the Seesaw model could have heavy leptons which couple exclusively to a specific flavor of lepton. Thus, the heavy lepton searched for in this analysis may only



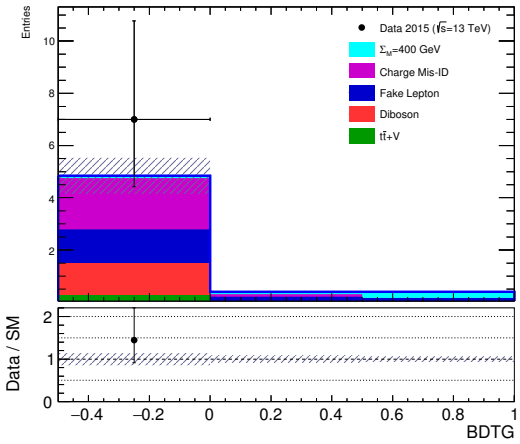
(a) 200 GeVmass point



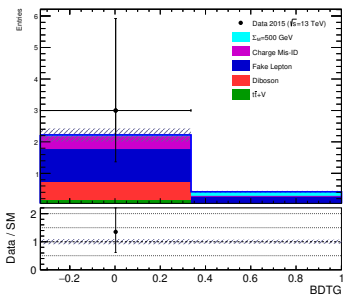
(b) 250 GeVmass point



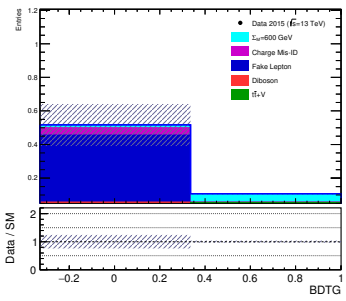
(c) 300 GeVmass point



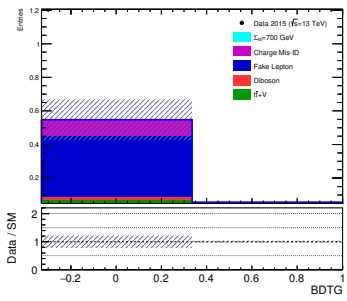
(d) 400 GeVmass point



(e) 500 GeVmass point



(f) 600 GeVmass point



(g) 700 GeVmass point

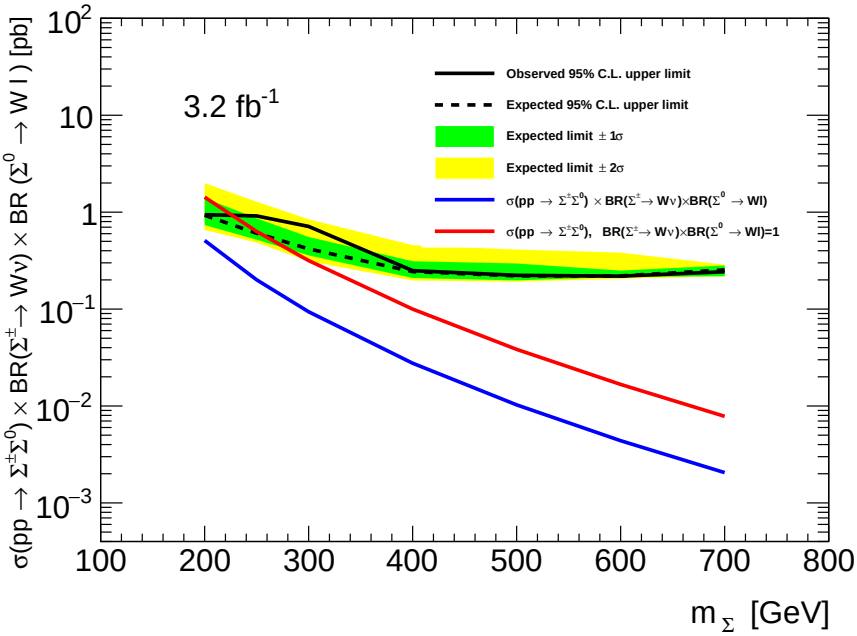
Figure 7.2: Data and MC events in the BDTG signal-like regions used for limit setting.

decay to electrons. This would increase the cross-section by roughly a factor of three, and correspondingly increase the sensitivity of the search. This was tested and the results are shown in Figure 7.3b. The blue line represents the theory predictions for the heavy lepton decaying to a W boson and a lepton. The red line assumes that the heavy lepton can only decay to a W boson and a lepton. With equal branching fractions to decay products no limit can be placed on the heavy neutrino mass. However, by constraining the heavy lepton to decay to W bosons and leptons, a limit is placed at 220 GeV/c^2. If the heavy lepton is further constrained to decay only to electron flavored leptons and W bosons, the heavy lepton mass is constrained to be greater than 420 GeV/c^2.

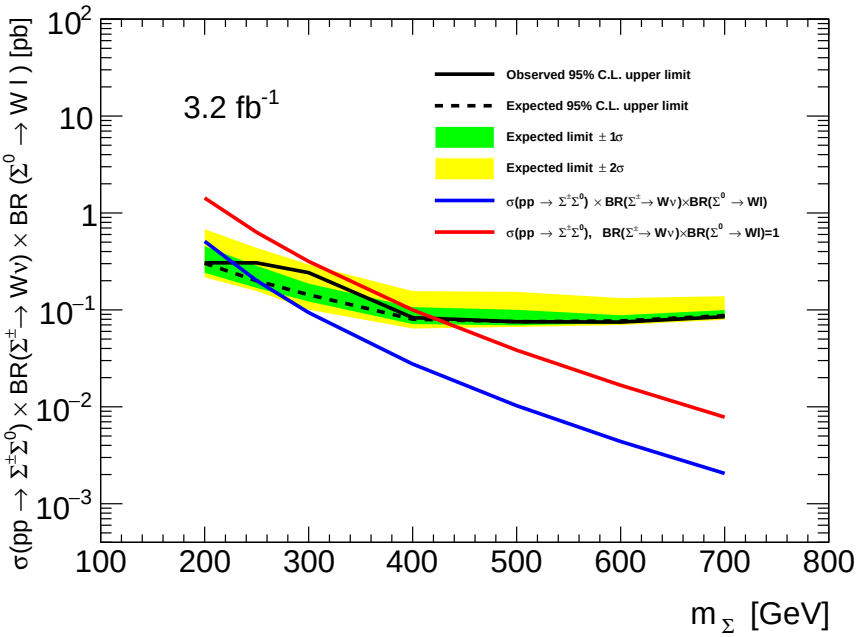
7.3.2 Outlook

The analysis focuses on comparing the techniques used in previous studies with a multivariate approach based on machine learning. There is significant improvement in sensitivity for the BDTG method, particularly at low heavy lepton mass. The BDTG method increases sensitivity by nearly a factor of three for the 200 GeV/c^2 heavy lepton mass. Figure 7.4 shows a comparison between the cut based method described in Section 6.4 and the BDTG method that is used for the main result.

The limits were extrapolated into the full end of Run 2 luminosity. It is projected that there will be roughly 100 fb^{-1} of data collected by the end of 2018. With this amount of data, the limits would be greatly improved, as is demonstrated in Figures 7.5. This shows the mass exclusion can be extended to over 400 GeV in the Type III Seesaw model with no decay constraints with the full Run 2 dataset.



(a) Limits at 3.2 fb^{-1}. The heavy lepton Σ is allowed to decay to all Standard Model leptons.



(b) Limits at 3.2 fb^{-1} with the heavy lepton Σ constrained to only decay to electrons and electron neutrinos.

Figure 7.3: Cross-section limits for 3.2 fb^{-1} of data. The black line shows the expected sensitivity of the BDTG method for 3.2 fb^{-1}. The green and yellow bands are 1 and 2 σ uncertainty ranges.

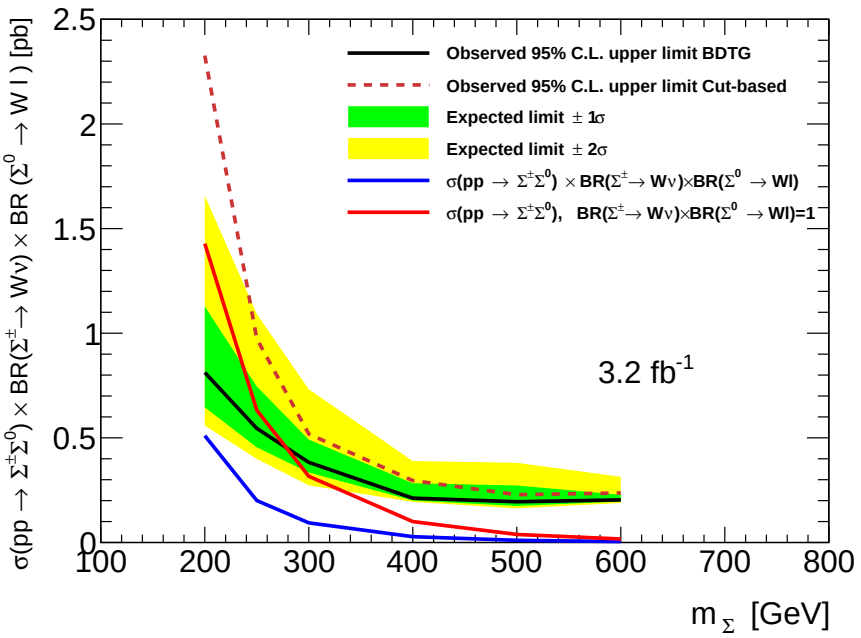
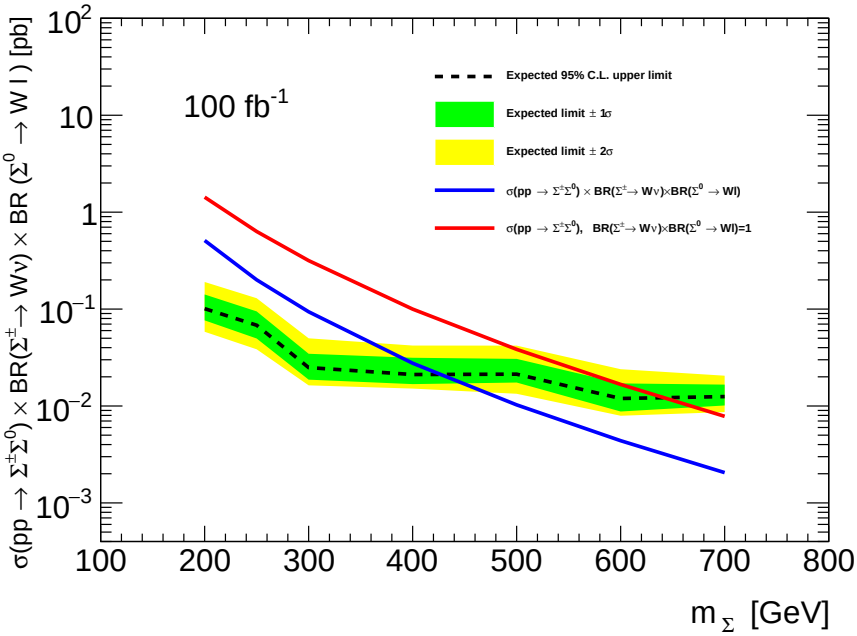
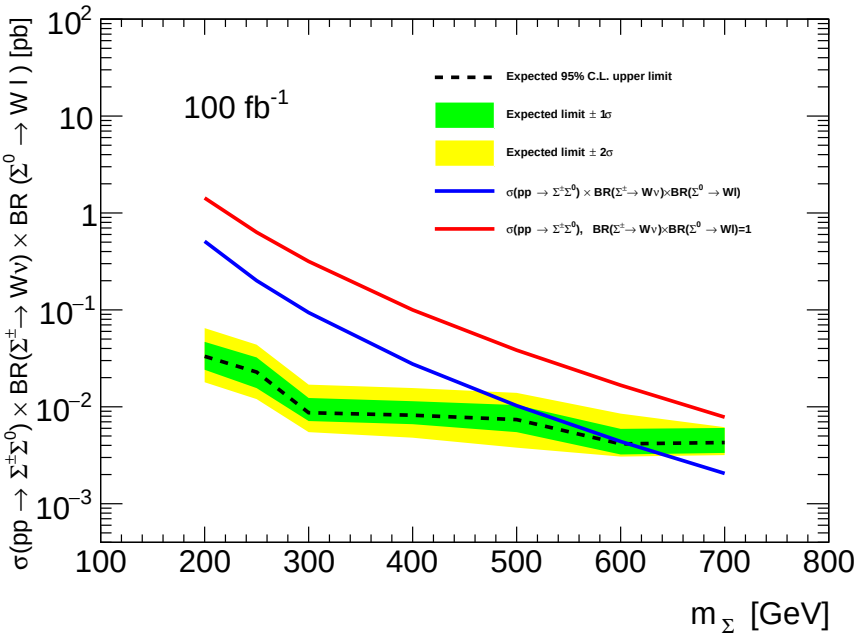


Figure 7.4: Cross-section limit comparison between cut-based and BDTG methods. The black line shows the expected sensitivity of the BDTG method for 3.2 fb^{-1}. The red dotted line shows the sensitivity for the cut based approach used in Run 1. The green and yellow bands are 1 and 2 σ uncertainty ranges. The blue line represents the theory predictions for the heavy lepton decaying to W plus lepton. The red line assumes that the heavy lepton can only decay to a W boson plus lepton.



(a) Limits at 100 fb^{-1}. The heavy lepton Σ is allowed to decay to all Standard Model leptons.



(b) Limits at 100 fb^{-1} with the heavy lepton Σ constrained to only decay to electrons and electron neutrinos.

Figure 7.5: Cross-section limits for 100 fb^{-1} of data. The black line shows the expected sensitivity of the BDTG method for 100 fb^{-1}. The green and yellow bands are 1 and 2 σ uncertainty ranges. The blue line represents the theory predictions for the heavy lepton decaying to W plus lepton. The red line assumes that the heavy lepton must decay to a W boson plus lepton.

CHAPTER 8. SUMMARY

This analysis performed a search for a heavy lepton predicted by the Type III Seesaw Mechanism. Data was collected via the ATLAS detector at CERN during the 2015 data collection period, which amounted to 3.2 fb^{-1} of integrated luminosity. Monte Carlo simulations were generated to estimate signal and background contributions and inform selection decisions.

The analysis examined the heavy lepton's decay to electrons specifically. In order to limit background contributions, the two electrons produced in the decay were required to have the same sign.

The focus of the analysis was to improve techniques in order to enhance sensitivity. Specifically, neural networks and boosted decision trees were trained and utilized to categorize data. The machine learning techniques were compared to an optimized set of kinematic selections used in a previous study. The machine learning techniques resulted in a three-fold increase in sensitivity at low mass, demonstrated in Figure 7.4

BIBLIOGRAPHY

[1] J. Barnes, *Early Greek Philosophy*. Classics Series. Penguin, 2001. https://books.google.com/books?id=ooq9SnHIdxIC.

[2] B. N. Menschutkin, *Early History of Mendeléeff's Periodic Law*, Nature **133** (1934) 946.

[3] J. Dalton, *On the absorption of gases by water and other liquids*, Philosophical Magazine **24** no. 93, (1806) 15–24, http://dx.doi.org/10.1080/14786440608563325, http://dx.doi.org/10.1080/14786440608563325.

[4] R. H. Petrucci and W. S. Harwood, *General chemistry: principles and modern applications*. Prentice Hall, 1997.

[5] J. Chadwick, *Possible Existence of a Neutron*, Nature **129** (1932) 312.

[6] M. Breidenbach et al., *Observed Behavior of Highly Inelastic Electron-Proton Scattering*, Phys. Rev. Lett. **23** (1969) 935–939.

[7] W. Pauli, *Open letter to the group of radioactive people at the Gauverein meeting in Tübingen*, 12, 1930. http://microboone-docdb.fnal.gov/cgi-bin/RetrieveFile?docid=953;filename=pauli%20letter1930.pdf.

[8] F. Reines and C. L. Cowan, *The neutrino*, Nature **178** (1956) 446–449.

[9] J. Schechter and J. W. F. Valle, *Neutrino Masses in SU(2) x U(1) Theories*, Phys. Rev. **D22** (1980) 2227.

[10] E. Ma and O. Popov, *Pathways to Naturally Small Dirac Neutrino Masses*, Phys. Lett. **B764** (2017) 142–144, arXiv:1609.02538 [hep-ph].

[11] M. Lindner, T. Ohlsson, and G. Seidl, *Seesaw mechanisms for Dirac and Majorana neutrino masses*, Phys. Rev. **D65** (2002) 053014, arXiv:hep-ph/0109264 [hep-ph].

[12] A. Zee, *A Theory of Lepton Number Violation, Neutrino Majorana Mass, and Oscillation*, Phys. Lett. **93B** (1980) 389, [Erratum: Phys. Lett.95B,461(1980)].

[13] R. N. Mohapatra and G. Senjanovic, *Neutrino Masses and Mixings in Gauge Models with Spontaneous Parity Violation*, Phys. Rev. **D23** (1981) 165.

[14] G. B. Gelmini and M. Roncadelli, *Left-Handed Neutrino Mass Scale and Spontaneously Broken Lepton Number*, Phys. Lett. **99B** (1981) 411–415.

[15] Y. Cai, X.-G. He, M. Ramsey-Musolf, and L.-H. Tsai, *Dark matter and radiative neutrino masses*, J. Phys. Conf. Ser. **384** (2012) 012026.

[16] R. N. Mohapatra and G. Senjanovic, *Neutrino Mass and Spontaneous Parity Violation*, Phys. Rev. Lett. **44** (1980) 912.

[17] A. Purcell, *Go on a particle quest at the first CERN webfest*, https://cdsweb.cern.ch/record/1473657/.

[18] Particle Data Group Collaboration, C. Patrignani et al., *Review of Particle Physics*, Chin. Phys. **C40** no. 10, (2016) 100001.

[19] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, *Detection of the free neutrino: A Confirmation*, Science **124** (1956) 103–104.

[20] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579–588.

[21] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266.

[22] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367–377.

[23] UA1 Collaboration, G. Arnison et al., *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at* $\sqrt{s} = 540$ GeV, Phys. Lett. **122B** (1983) 103–116, [,611(1983)].

[24] UA1 Collaboration, G. Arnison et al., *Experimental Observation of Lepton Pairs of Invariant Mass Around* 95 GeV/c^2 *at the CERN SPS Collider*, Phys. Lett. **126B** (1983) 398–410.

[25] UA2 Collaboration, P. Bagnaia et al., *Evidence for* $Z^0 \to e^+e^-$ *at the CERN anti-p p Collider*, Phys. Lett. **129B** (1983) 130–140.

[26] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509.

[27] *Higgs boson theory and phenomenology*, Prog. Part. Nucl. Phys. **50** (2003) 63–152, arXiv:hep-ph/0208209 [hep-ph].

[28] J. F. Gunion and H. E. Haber, *The CP conserving two Higgs doublet model: The Approach to the decoupling limit*, Phys. Rev. **D67** (2003) 075019, arXiv:hep-ph/0207010 [hep-ph].

[29] G. Panico and A. Wulzer, *The Composite Nambu-Goldstone Higgs*, Lect. Notes Phys. **913** (2016) pp.1–316, arXiv:1506.01961 [hep-ph].

[30] ATLAS Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[31] CMS Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of* 125 GeV *with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

[32] M. Goldhaber, L. Grodzins, and A. W. Sunyar, *Helicity of Neutrinos*, Phys. Rev. **109** (1958) 1015–1017.

[33] B. Pontecorvo, *Inverse beta processes and nonconservation of lepton charge*, Sov. Phys. JETP **7** (1958) 172–173, [Zh. Eksp. Teor. Fiz.34,247(1957)].

[34] S. Turck-Chize, *The Standard Solar Model and beyond*, Journal of Physics: Conference Series **665** no. 1, (2016) 012078, http://stacks.iop.org/1742-6596/665/i=1/a=012078.

[35] K. Lodders, *Solar System Abundances and Condensation Temperatures of the Elements*, Astrophys. J. **591** (2003) 1220–1247.

[36] B. T. Cleveland, T. Daily, R. Davis, Jr., J. R. Distel, K. Lande, C. K. Lee, P. S. Wildenhain, and J. Ullman, *Measurement of the solar electron neutrino flux with the Homestake chlorine detector*, Astrophys. J. **496** (1998) 505–526.

[37] J. N. Bahcall and R. K. Ulrich, *Solar Models, Neutrino Experiments and Helioseismology*, Rev. Mod. Phys. **60** (1988) 297–372.

[38] M. Fukugita and T. Yanagida, *Physics of neutrinos and applications to astrophysics*, http://public.eblib.com/choice/PublicFullRecord.aspx?p=3099134.

[39] SAGE Collaboration, J. N. Abdurashitov et al., *Measurement of the solar neutrino capture rate with gallium metal. III: Results for the 2002–2007 data-taking period*, Phys. Rev. **C80** (2009) 015807, arXiv:0901.2200 [nucl-ex].

[40] K2K Collaboration, M. H. Ahn et al., *Measurement of Neutrino Oscillation by the K2K Experiment*, Phys. Rev. **D74** (2006) 072003, arXiv:hep-ex/0606032 [hep-ex].

[41] MINOS Collaboration, D. G. Michael et al., *Observation of muon neutrino disappearance with the MINOS detectors and the NuMI neutrino beam*, Phys. Rev. Lett. **97** (2006) 191801, arXiv:hep-ex/0607088 [hep-ex].

[42] F. Boehm et al., *Final results from the Palo Verde neutrino oscillation experiment*, Phys. Rev. **D64** (2001) 112001, arXiv:hep-ex/0107009 [hep-ex].

[43] H. Murayama, *Neutrino Physics*, http://hitoshi.berkeley.edu/neutrino/.

[44] G. C. Branco, R. G. Felipe, and F. R. Joaquim, *Leptonic CP Violation*, Rev. Mod. Phys. **84** (2012) 515–565, arXiv:1111.5332 [hep-ph].

[45] Troitsk Collaboration, V. N. Aseev et al., *An upper limit on electron antineutrino mass from Troitsk experiment*, Phys. Rev. **D84** (2011) 112003, arXiv:1108.5034 [hep-ex].

[46] KATRIN Collaboration, T. Thümmler, *Introduction to direct neutrino mass measurements and KATRIN*, Nucl. Phys. Proc. Suppl. **229-232** (2012) 146–151, arXiv:1012.2282 [hep-ex].

[47] KATRIN Collaboration, S. Mertens, *Status of the KATRIN Experiment and Prospects to Search for keV-mass Sterile Neutrinos in Tritium β-decay*, Phys. Procedia **61** (2015) 267–273.

[48] S. Mertens, T. Lasserre, S. Groh, G. Drexlin, F. Glueck, A. Huber, A. W. P. Poon, M. Steidl, N. Steinbrink, and C. Weinheimer, *Sensitivity of Next-Generation Tritium Beta-Decay Experiments for keV-Scale Sterile Neutrinos*, JCAP **1502** no. 02, (2015) 020, arXiv:1409.0920 [physics.ins-det].

[49] KATRIN Collaboration, J. Angrik et al., *KATRIN design report 2004*,.

[50] L. M. Krauss, *DARK MATTER IN THE UNIVERSE .... AND IN THE LABORATORY*, pp. , 665–678. 1987. https://inspirehep.net/record/247981/files/Pages_from_C87-03-08_665-678.pdf.

[51] L. Bergstrm, *Nonbaryonic dark matter: Observational evidence and detection methods*, Rept. Prog. Phys. **63** (2000) 793, arXiv:hep-ph/0002126 [hep-ph].

[52] G. Bertone, D. Hooper, and J. Silk, *Particle dark matter: Evidence, candidates and constraints*, Phys. Rept. **405** (2005) 279–390, arXiv:hep-ph/0404175 [hep-ph].

[53] A. Albert, *Searching for Dark Matter with Cosmic Gamma Rays*. IOP Publishing, 2016. https://doi.org/10.1088%2F978-1-6817-4269-4.

[54] P. A. M. Dirac, *A Theory of Electrons and Protons*, Proc. Roy. Soc. Lond. **A126** (1930) 360.

[55] A. D. Sakharov, *Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32–35, [Usp. Fiz. Nauk161,61(1991)].

[56] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the* 2π *Decay of the* K_2^0 *Meson*, Phys. Rev. Lett. **13** (1964) 138–140.

[57] A. Riotto, *Theories of baryogenesis*, pp. , 326–436. 1998. arXiv:hep-ph/9807454 [hep-ph].
https://inspirehep.net/record/473645/files/arXiv:hep-ph_9807454.pdf.

[58] S. M. Bilenky, *Neutrino in Standard Model and beyond*, Phys. Part. Nucl. **46** no. 4, (2015) 475–496, arXiv:1501.00232 [hep-ph].

[59] M. Thomson, *Modern particle physics*. Cambridge University Press, New York, 2013. http://www-spires.fnal.gov/spires/find/books/www?cl=QC793.2.T46::2013.

[60] H. Murayama, *The origin of neutrino mass*, Physics World **15** no. 5, (2002) 35, http://stacks.iop.org/2058-7058/15/i=5/a=36.

[61] M. Thomson, *Modern particle physics*. Cambridge University Press, 2014.

[62] F. Ferroni, F. Vissani, and C. Brofferio, *Proceedings, International School of Physics "Enrico Fermi", 170th Course, "Measurements of Neutrino Mass"*, vol. 170. 2009.

[63] W. Grimus, L. Lavoura, and B. Radovcic, *Type II seesaw mechanism for Higgs doublets and the scale of new physics*, Phys. Lett. **B674** (2009) 117–121, arXiv:0902.2325 [hep-ph].

[64] P. Fileviez Perez, *Type III Seesaw and Left-Right Symmetry*, JHEP **03** (2009) 142, arXiv:0809.1202 [hep-ph].

[65] R. Franceschini, T. Hambye, and A. Strumia, *Type-III see-saw at LHC*, Phys. Rev. **D78** (2008) 033002, arXiv:0805.1613 [hep-ph].

[66] E. Ma, *Pathways to naturally small neutrino masses*, Phys. Rev. Lett. **81** (1998) 1171–1174, arXiv:hep-ph/9805219 [hep-ph].

[67] B. Bajc and G. Senjanovic, *Seesaw at LHC*, JHEP **08** (2007) 014, arXiv:hep-ph/0612029 [hep-ph].

[68] P. Fileviez Perez, *Renormalizable adjoint SU(5)*, Phys. Lett. **B654** (2007) 189–193, arXiv:hep-ph/0702287 [hep-ph].

[69] M. K. Parida and A. Raychaudhuri, *Inverse see-saw, leptogenesis, observable proton decay and* $\Delta_{\rm R}^{\pm\pm}$ *in SUSY SO(10) with heavy* W_R, Phys. Rev. **D82** (2010) 093017, arXiv:1007.5085 [hep-ph].

[70] R. Foot, H. Lew, X. G. He, and G. C. Joshi, *Seesaw Neutrino Masses Induced by a Triplet of Leptons*, Z. Phys. **C44** (1989) 441.

[71] C. Biggio and F. Bonnet, *Implementation of the Type III Seesaw Model in FeynRules/MadGraph and Prospects for Discovery with Early LHC Data*, Eur. Phys. J. **C72** (2012) 1899, arXiv:1107.3463 [hep-ph].

[72] L3 Collaboration, P. Achard et al., *Search for heavy neutral and charged leptons in e^+e^- annihilation at LEP*, Phys. Lett. **B517** (2001) 75–85, arXiv:hep-ex/0107015 [hep-ex].

[73] CMS Collaboration, S. Chatrchyan et al., *Search for heavy lepton partners of neutrinos in proton-proton collisions in the context of the type III seesaw mechanism*, Phys. Lett. **B718** (2012) 348–368, arXiv:1210.1797 [hep-ex].

[74] ATLAS Collaboration, G. Aad et al., *Search for heavy lepton resonances decaying to a Z boson and a lepton in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **09** (2015) 108, arXiv:1506.01291 [hep-ex].

[75] ATLAS Collaboration, G. Aad et al., *Search for type-III Seesaw heavy leptons in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*, Phys. Rev. **D92** no. 3, (2015) 032001, arXiv:1506.01839 [hep-ex].

[76] GERDA Collaboration, M. Agostini et al., *Results on Neutrinoless Double-β Decay of ^{76}Ge from Phase I of the GERDA Experiment*, Phys. Rev. Lett. **111** no. 12, (2013) 122503, arXiv:1307.4720 [nucl-ex].

[77] Gerda Collaboration, M. Miloradovic, *The GERDA Experiment: Search for the Neutrinoless Double Beta Decay*, pp. , 41–46. 2016. https://inspirehep.net/record/1501314/files/fulltext.pdf.

[78] M. Fukugita and T. Yanagida, *Baryogenesis Without Grand Unification*, Phys. Lett. **B174** (1986) 45–47.

[79] ATLAS Collaboration, M. Stockton, *The ATLAS level-1 central trigger*, J. Phys. Conf. Ser. **331** (2011) 022041.

[80] LHC Higgs Cross Section Working Group Collaboration, J. R. Andersen et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, arXiv:1307.1347 [hep-ph].

[81] ATLAS Collaboration, M. Aaboud et al., *Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C76** no. 12, (2016) 653, arXiv:1608.03953 [hep-ex].

[82] ATLAS Collaboration, A. Koutoulaki, *The ITk Strip Tracker for the phase-II upgrade of the ATLAS detector of the HL-LHC*, JINST **12** no. 04, (2017) C04022.

[83] *Distributing and storing data efficiently by means of special datasets in the ATLAS collaboration*, vol. 331. 2011.

[84] J. Steinberger, *Learning about particles: 50 privileged years*. 2005.

[85] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph].

[86] ATLAS Collaboration, G. Aad et al., *ATLAS magnet system: Technical design report*,.

[87] ATLAS Collaboration, G. Aad et al., *ATLAS muon spectrometer: Technical design report*. 1997.

[88] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[89] ATLAS Collaboration, C. Lapoire, *Operational experience with the ATLAS Pixel Detector at the LHC*, Nucl. Instrum. Meth. **A699** (2013) 149–152.

[90] A. La Rosa, *The ATLAS Insertable B-Layer: from construction to operation*, JINST **11** no. 12, (2016) C12036, arXiv:1610.01994 [physics.ins-det].

[91] L. Rossi, P. Fischer, T. Rohe, and N. Wermes, *Pixel Detectors*. Particle Acceleration and Detection. Springer-Verlag, Berlin, 2006. http://www.springer.com/book/3-540-28332-3.

[92] ATLAS Collaboration, G. Aad et al., *ATLAS inner detector: Technical design report. Vol. 1*. 1997.

[93] ATLAS LArG Collaboration, H. Wilkens, *The ATLAS liquid argon calorimeter: An overview*, J. Phys. Conf. Ser. **160** (2009) 012043.

[94] ATLAS Collaboration, A. M. Henriques Correia, *The ATLAS Tile Calorimeter*, Tech. Rep. ATL-TILECAL-PROC-2015-002, CERN, Geneva, Mar, 2015. https://cds.cern.ch/record/2004868.

[95] ATLAS Collaboration, G. Aad et al., *Commissioning of the ATLAS Muon Spectrometer with Cosmic Rays*, Eur. Phys. J. **C70** (2010) 875–916, arXiv:1006.4384 [physics.ins-det].

[96] A. Pich, *Precision Tau Physics*, Prog. Part. Nucl. Phys. **75** (2014) 41–85, arXiv:1310.7922 [hep-ph].

[97] ATLAS Collaboration, S. J. Thais, *Electron ID in ATLAS Run 2*, https://cds.cern.ch/record/2265356.

[98] J. Alison, K. Brendlinger, S. Heim, J. Kroll, and C. M. Lester, *Description and Performance of the Electron Likelihood Tool at ATLAS using 2012 LHC Data*, Tech. Rep. ATL-COM-PHYS-2013-378, CERN, Geneva, Apr, 2013. https://cds.cern.ch/record/1537410.

[99] *Isolation Selection Tool*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/IsolationSelectionTool.

[100] *Muon Selection Tool*, https://twiki.cern.ch/twiki/bin/view/Atlas/MuonSelectionTool.

[101] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[102] ATLAS Collaboration, G. Aad et al., *Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC*, Eur. Phys. J. **C71** (2011) 1630, arXiv:1101.2185 [hep-ex].

[103] ATLAS Collaboration, G. Aad et al., *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. **C76** no. 11, (2016) 581, arXiv:1510.03823 [hep-ex].

[104] *Tagging and suppression of pileup jets with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2014-018, CERN, Geneva, May, 2014. https://cds.cern.ch/record/1700870.

[105] *Pile-up subtraction and suppression for jets in ATLAS*, Tech. Rep. ATLAS-CONF-2013-083, CERN, Geneva, Aug, 2013. https://cds.cern.ch/record/1570994.

[106] *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV*, Tech. Rep. ATL-PHYS-PUB-2015-015, CERN, Geneva, Jul, 2015. https://cds.cern.ch/record/2037613.

[107] ATLAS Collaboration,, *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*, May, 2015. https://cds.cern.ch/record/2016323.

[108] *Jet Vertex Tagger*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetVertexTagger.

[109] *How To Clean Jets 2015*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/HowToCleanJets2015.

[110] *B Hadron Identification*. 2017. https://www-d0.fnal.gov/Run2Physics/top/singletop_observation/b_tagging_graphic.png.

[111] ATLAS Collaboration, G. Aad et al., *Performance of b-Jet Identification in the ATLAS Experiment*, JINST **11** no. 04, (2016) P04008, arXiv:1512.01094 [hep-ex].

[112] ATLAS Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].

[113] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303.

[114] M. H. Seymour and M. Marx, *Monte Carlo Event Generators*, pp. , 287–319. 2013. arXiv:1304.6677 [hep-ph]. https://inspirehep.net/record/1229804/files/arXiv:1304.6677.pdf.

[115] *ATLAS Generators and Interface Documentation*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/McGeneratorsForAtlas.

[116] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[117] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv:hep-ph/0409146 [hep-ph].

[118] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001, arXiv:hep-ph/0206293 [hep-ph].

[119] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv:0811.4622 [hep-ph].

[120] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv:1405.0301 [hep-ph].

[121] P. M. Nadolsky, *Theory of W and Z boson production*, AIP Conf. Proc. **753** (2005) 158–170, arXiv:hep-ph/0412146 [hep-ph], [,158(2004)].

[122] ATLAS Collaboration, G. Aad et al., *Measurement of $W^{\pm}$ and Z-boson production cross sections in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. **B759** (2016) 601–621, arXiv:1603.09222 [hep-ex].

[123] ATLAS Collaboration, M. Aaboud et al., *Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s}$=13 TeV with the ATLAS detector*, Phys. Lett. **B761** (2016) 136–157, arXiv:1606.02699 [hep-ex].

[124] F. Kohn, *Measurement of the charge asymmetry in top quark pair production in pp collision data at $\sqrt{s} = 7$ TeV using the ATLAS detector*. PhD thesis, U. Gottingen (main), 2012. arXiv:1204.0952 [hep-ex]. https://inspirehep.net/record/1103034/files/arXiv:1204.0952.pdf.

[125] *B Hadron Identification*. 2017. http://www-d0.fnal.gov/Run2Physics/top/top_public_web_pages/feynman_diagrams/feynman_ttbar_ljets_beamline.eps.

[126] ATLAS Collaboration, M. Aaboud et al., *Measurement of the inclusive cross-sections of single top-quark and top-antiquark t-channel production in pp*

collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector, JHEP **04** (2017) 086, arXiv:1609.03920 [hep-ex].

[127] ATLAS Collaboration, M. Aaboud et al., *Measurements of top-quark pair to Z-boson cross-section ratios at $\sqrt{s} = 13, 8, 7$ TeV with the ATLAS detector*, JHEP **02** (2017) 117, arXiv:1612.03636 [hep-ex].

[128] ATLAS Collaboration, M. Aaboud et al., *Measurement of the $W^{\pm}Z$ boson pair-production cross section in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*, Phys. Lett. **B762** (2016) 1–22, arXiv:1606.04017 [hep-ex].

[129] ATLAS Collaboration, M. Aaboud et al., *Search for triboson $W^{\pm}W^{\pm}W^{\mp}$ production in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Eur. Phys. J. **C77** no. 3, (2017) 141, arXiv:1610.05088 [hep-ex].

[130] *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ production cross sections in multilepton final states using 3.2 fb^{-1} of pp collisions at 13 TeV at the LHC*, Tech. Rep. ATLAS-CONF-2016-003, CERN, Geneva, Mar, 2016. https://cds.cern.ch/record/2138947.

[131] CMS Collaboration, S. Chatrchyan et al., *Measurement of associated production of vector bosons and top quark-antiquark pairs at sqrt(s) = 7 TeV*, Phys. Rev. Lett. **110** (2013) 172002, arXiv:1303.3239 [hep-ex].

[132] S. D. Drell and T.-M. Yan, *Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies*, Phys. Rev. Lett. **25** (1970) 316–320, [Erratum: Phys. Rev. Lett.25,902(1970)].

[133] G. Siragusa, V. Pleskot, R. Pottgen, N. Benekos, G. Zacharis, I. Panagoulias, D. Milstead, S. Raimund, and T. Papadopoulou, *Search for First and Second generation Scalar Leptoquarks at sqrt(s)=13 TeV with the ATLAS detector*, Tech. Rep. ATL-COM-PHYS-2015-1210, CERN, Geneva, Sep, 2015.

https://cds.cern.ch/record/2055510. Conversion of the 2015 1st and 2nd generation leptoquark search supporting documentation into an INT note. Members of the EdBoard (Borut Kersevan, Masahiro Kuze and Peter Kluit) agree.

[134] P. O. DeViveiros, S. Dube, M. Hance, B. Heinemann, I. Hinchliffe, M. Hurwitz, O. Igonkina, P. Jackson, J. Mahlstedt, N. Soni, and D. Yu, *Search for New Physics in Events with Three Charged Leptons*, Tech. Rep. ATL-COM-PHYS-2013-651, CERN, Geneva, May, 2013. https://cds.cern.ch/record/1547859.

[135] M. Muskinja, B. P. Kersevan, M. Mikuz, and A. Gorisek, *Search for high-mass new phenomena in the same-sign di-electron final state using proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Tech. Rep. ATL-COM-PHYS-2016-484, CERN, Geneva, May, 2016. https://cds.cern.ch/record/2151853.

[136] D. J. Summers, *Transverse mass as a means of measuring the W width at the Tevatron*, Phys. Lett. **B392** (1997) 216–222, arXiv:hep-ph/9608332 [hep-ph].

[137] M. Muskinja et al., *Search for doubly-charged Higgs boson production in multi-lepton final states with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV*, Tech. Rep. ATL-COM-PHYS-2016-1415, CERN, Geneva, Sep, 2016. https://cds.cern.ch/record/2218092. Uploading the 2016 ICHEP supporting note as a place holder for Moriond 2017 note.

[138] ATLAS Collaboration, G. Aad et al., *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur. Phys. J. **C75** (2015) 17, arXiv:1406.0076 [hep-ex].

[139] G. Cowan and E. Gross, *Discovery significance with statistical uncertainty in the background estimate*, p. , 213. 5, 2008.

[140] J. R. Quinlan, *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993.

[141] L. Breiman et al., *Classification and regression trees*. Chapman & Hall, New York, N.Y., 1984.

[142] R. E. Schapire and Y. Freund, *Boosting: Foundations and Algorithms*. The MIT Press, 2012.

[143] T. Hastie, *The Elements of Statistical Learning : Data mining, Inference, and Prediction*. Springer, New York, 2nd ed., 2009.

[144] *General regression and over fitting*, Sep, 2013. `https://shapeofdata.wordpress.com/2013/03/26/general-regression-and-over-fitting/`.

[145] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, `arXiv:physics/0703039`.

[146] M. Hazewinkel, V. I. Matveevic, V. I. Bityutskov, R. V. Gamkrelidze, and P. U. Vasilevic, *Encyclopaedia of mathematics: an updated and annotated translation of the Soviet Mathematical Encyclopaedia*. Reidel imprint of Kluwer Academic publ, 1988.

[147] *Good Runs List For Analysis Run 2*, `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/GoodRunListsForAnalysisRun2`.

[148] ATLAS Collaboration,, *Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at* $\sqrt{s} = 13$ TeV, Tech. Rep. ATL-PHYS-PUB-2016-015, CERN, Geneva, Aug, 2016. `https://cds.cern.ch/record/2203514`.

[149] ATLAS Collaboration,, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, Tech. Rep. ATLAS-CONF-2016-024, CERN, Geneva, Jun, 2016. `https://cds.cern.ch/record/2157687`.

[150] J.-B. De Vivie De Regie and others., *ATLAS electron, photon and muon isolation in Run 2*, Tech. Rep. ATL-COM-PHYS-2017-290, CERN, Geneva, Mar, 2017. https://cds.cern.ch/record/2256658. This note contains the Moriond 2017 recommendations. It will be updated when new recommendations become available.

[151] *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at* $\sqrt{s}$*= 13 TeV*, Tech. Rep. ATL-PHYS-PUB-2015-027, CERN, Geneva, Jul, 2015. https://cds.cern.ch/record/2037904.

[152] ATLAS Collaboration, M. Aaboud et al., *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, arXiv:1703.09665 [hep-ex].

[153] M. Baak, G. J. Besjes, D. Cte, A. Koutsman, J. Lorenz, and D. Short, *HistFitter software framework for statistical data analysis*, Eur. Phys. J. **C75** (2015) 153, arXiv:1410.1280 [hep-ex].

[154] L. Moneta, K. Belasco, K. S. Cranmer, S. Kreiss, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke, and M. Wolf, *The RooStats Project*, PoS **ACAT2010** (2010) 057, arXiv:1009.1003 [physics.data-an].

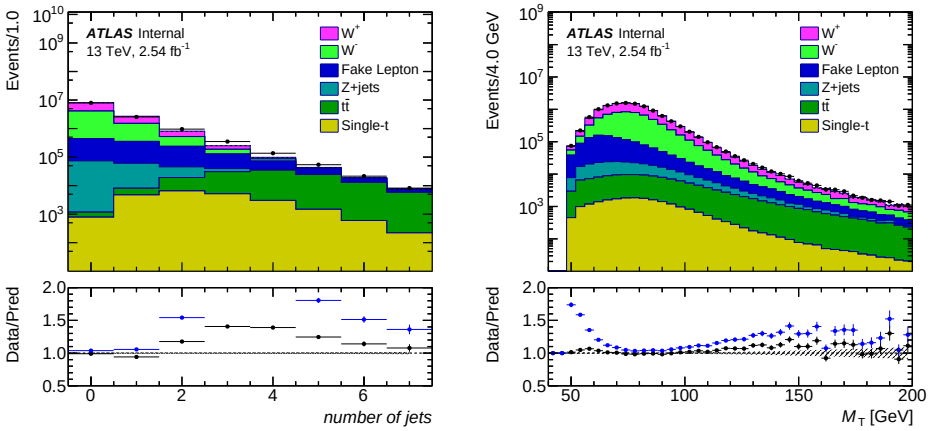
APPENDIX A. FAKE FACTOR VALIDATION IN THE W BOSON ENRICHED REGION

The W boson normalization matches well at low jet multiplicities, and the fake factor method has been previously validated in that region with the same MC samples [135]. Figures A.1 the kinematic distributions in the W boson enriched region, comparing data with prompt electrons from MC. The blue dots below each figure show this comparison. The black dots take into account the Fake Lepton contributions as well. With the inclusion of the Fake Lepton sample, there is good agreement between data and predictions. However, there is a noticeable discrepancy most evident in Figure A.1a where there is more than one jet. As the two first bins are several orders of magnitude larger than the additional jet bins, the discrepancy is not obvious when examining other variables.

Figure A.2 shows the same W boson enriched region, but with an additional cut requiring each event to have at least two jets. There is a large discrepancy of $\approx 50\%$ without accounting for the fake leptons. With the fake leptons there is still a missing 20% at least. Figure A.2b makes it clear through the shape of the backgrounds that the discrepancy cannot be due to the Fake Lepton background alone. Rather, the missing shape is consistent with the shape of the W boson backgrounds.

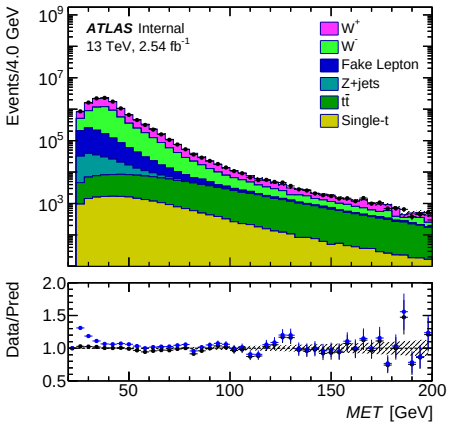
Using a fit with a free parameter on the normalization of the W boson backgrounds, the total background was matched to the data. The fit give an adjustment of 1.35 in the overall normalization of the W backgrounds for events with two or more jets.

Figure A.3 shows the same distributions as Figure A.1 with the 35% correction applied. The discrepancy related to jet multiplicity is greatly reduced and better agreement between data and predictions is demonstrated.

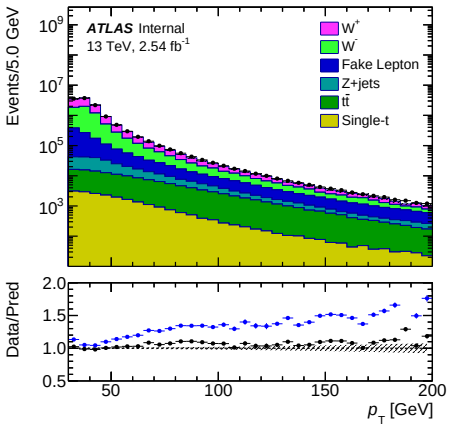


(a) Jet Multiplicity

(b) Transverse mass

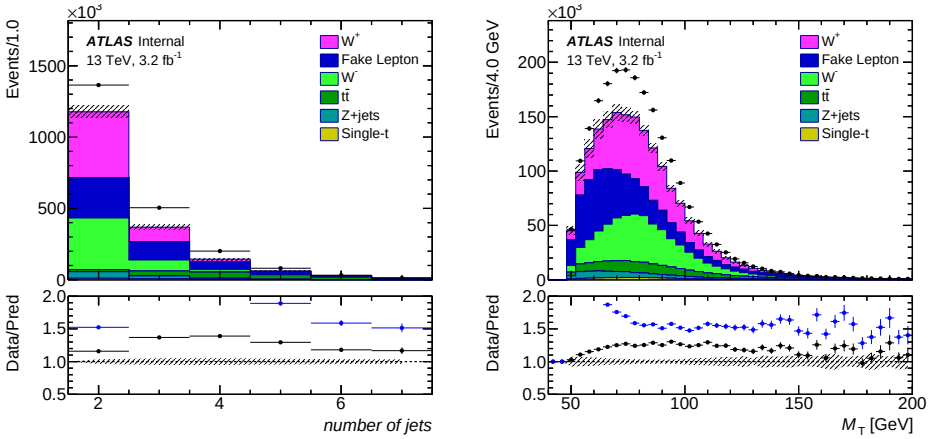


(c) $E_{\mathrm{T}}^{\mathrm{miss}}$



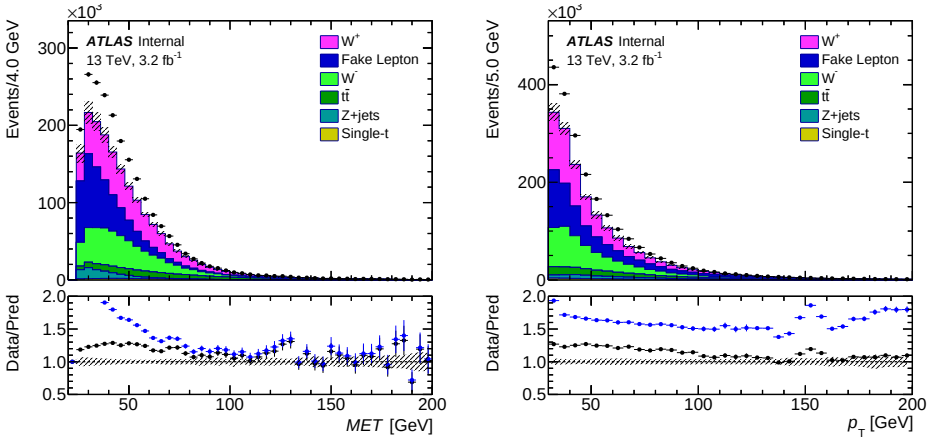
(d) Leading electron p_{T}

Figure A.1: W enriched region where no jet cut has been applied. Good agreement can be seen in all kinematic variables except there is a marked discrepancy where there are 2-6 jets in the event



(a) Jet Multiplicity

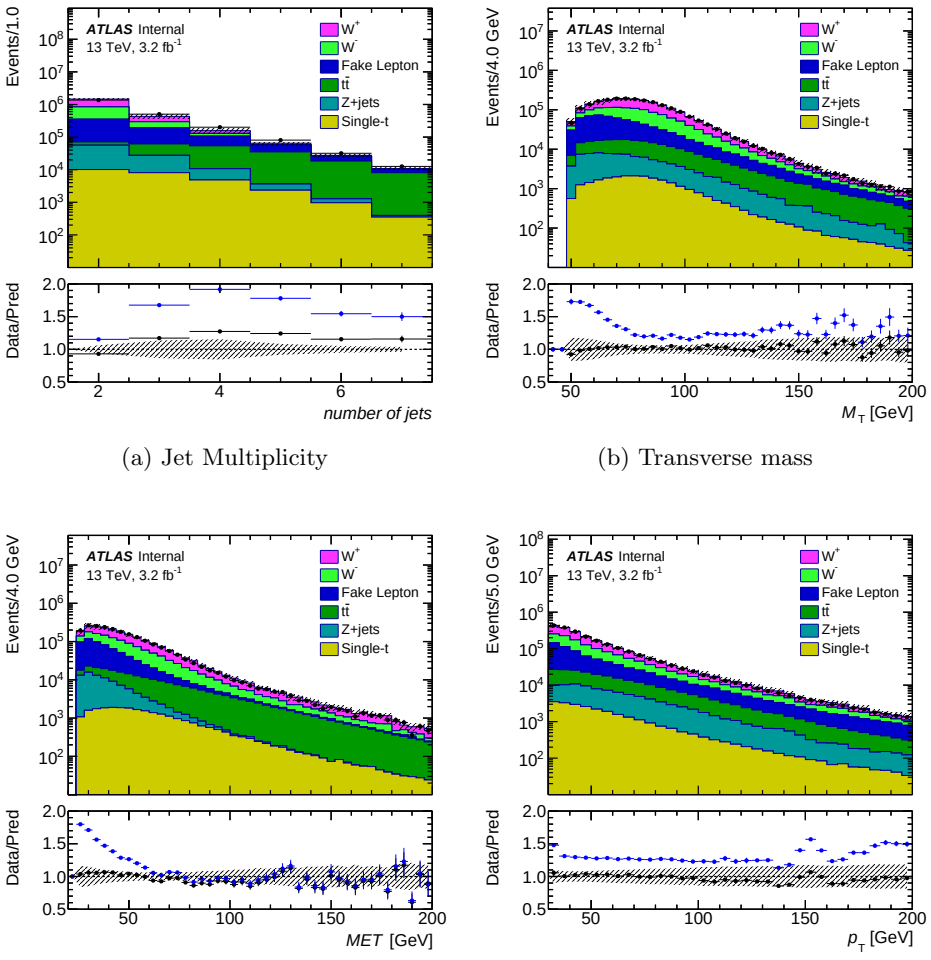
(b) Transverse mass



(c) $E_{\mathrm{T}}^{\mathrm{miss}}$

(d) Leading electron p_{T}

Figure A.2: W enriched region with 2+ jets with no correcting normalization



(a) Jet Multiplicity

(b) Transverse mass

(c) $E_{\mathrm{T}}^{\mathrm{miss}}$

(d) Leading electron p_{T}

Figure A.3: W enriched region with 2+ jets after the correcting normalization has been applied to the W boson samples

APPENDIX B. FAKE ELECTRON VALIDATION

The same-sign only fake lepton sample is quite statistically limited. This is not a problem with simple analysis techniques, such as a cut-and-count method. There is more of an issue however when using multivariate techniques which is the technique being tested in this analysis. Figure B.1 shows some control region plots demonstrating good agreement between data and background samples.

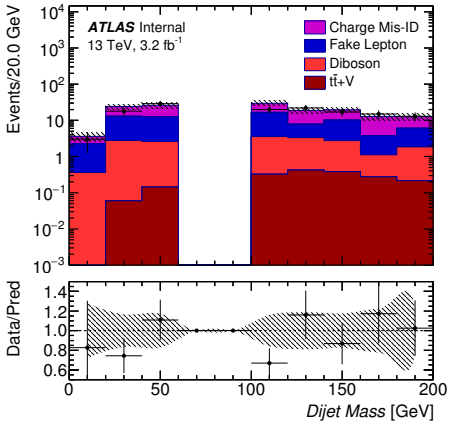
The same-sign and opposite-sign fake contributions are expected to behave very similarly. The objects which are reconstructed as jets are either uncharged, as with the case of photons or gluons, or originate from quarks and can carry some charge information. However in the uncharged case, there can be no difference between same-sign and opposite-sign events as the reconstruction algorithm for the electron has selected a charge for the object. The choice should be entirely random.

The quarks however do have a notable difference in rates between the number of same-sign and opposite-sign fakes. Jet charge does correlate with lepton charge but with a wide error. In other words there is a wide mixing of quark and lepton pairs with different signs into same-sign regions and vis-versa. Ultimately the same-sign fake lepton contributions are not fundamentally different from opposite-sign contributions. To prove this definitively is very difficult given the low statistics available. The low statistics is also a significant problem when dealing with machine learning, which as a system, typically relies on a large training sample.

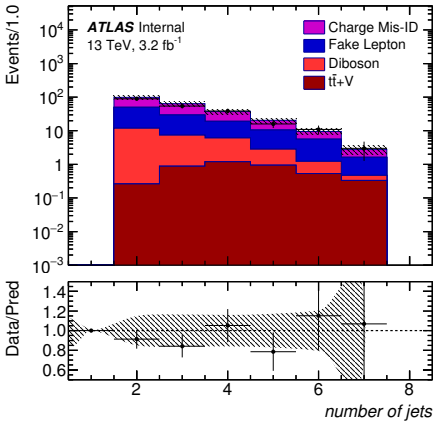
As a solution to the statistical limitation, it was decided to not make the same-sign selection, and instead apply a normalization on the inclusive fake electron contribution to the scale of the same-sign only contribution.

The normalization factor was found to be 0.21. Figure B.2 shows the inclusive and normalized sample. It is consistent with the same-sign only selection, but with ≈ 5 times the statistics, thus making it much more suitable for use with the multivariate techniques. As the kinematic behavior is consistent with the same-sign only fake sample, it was deemed that the normalized sample is an appropriate solution to the statistical issues.

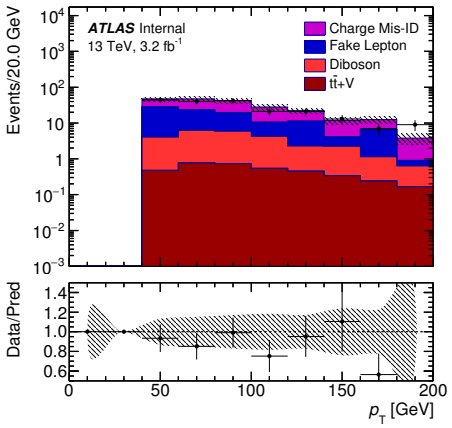
Ideally this method would not be used, the sample which is important are events which are read by the detector as same-sign. When this analysis is repeated for the full Run 2 dataset, this method should no longer be necessary as the statistics will increase by roughly a factor of 30.



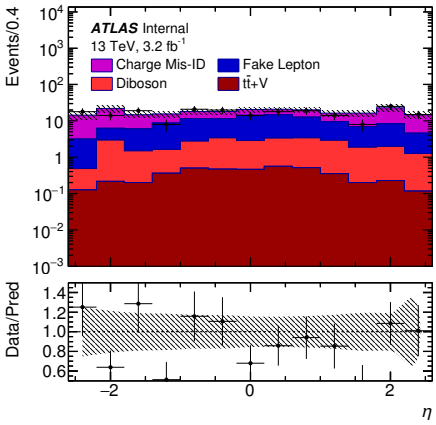
(a) Invariant DiJet Mass



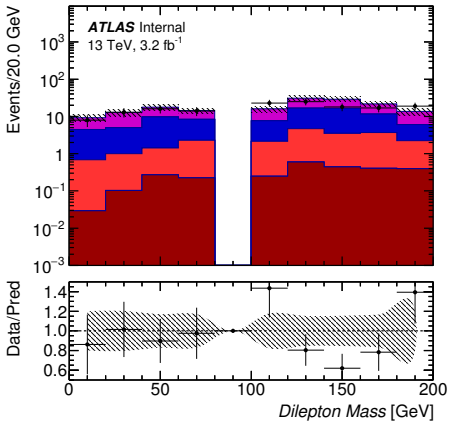
(b) Jet Multiplicity



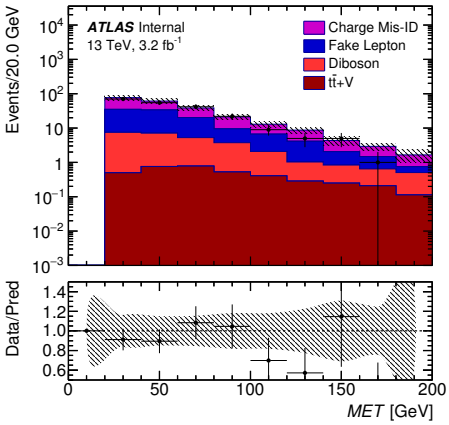
(c) Leading electron p_{T}



(d) Leading electron η

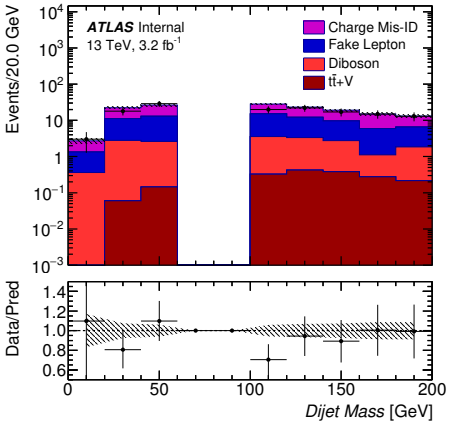


(e) Di-Electron Mass

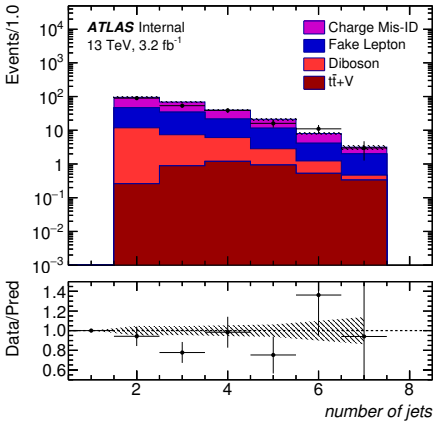


(f) $E_{\mathrm{T}}^{\mathrm{miss}}$

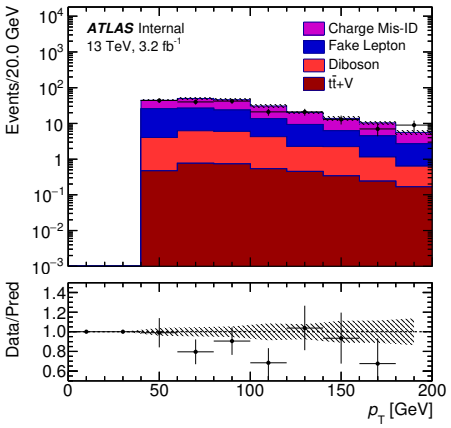
Figure B.1: W sidebands region, with same-sign electron selection. Additionally, there is a $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV and leading electron $p_{\mathrm{T}} > 40$ GeV cut for cleaning. The sidebands are defined as having a di-jet mass > 100 GeV or < 60 GeV. Two electrons are required in order to make the same-sign selection. This region is useful for examining the contribution of fake electrons.



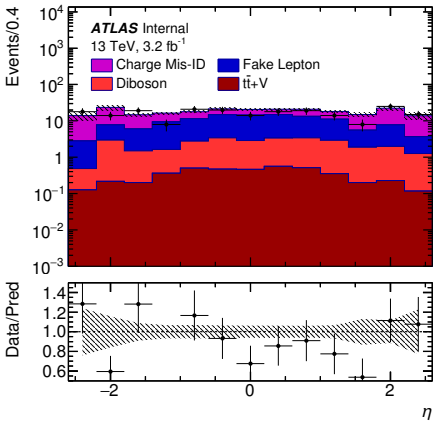
(a) Invariant DiJet Mass



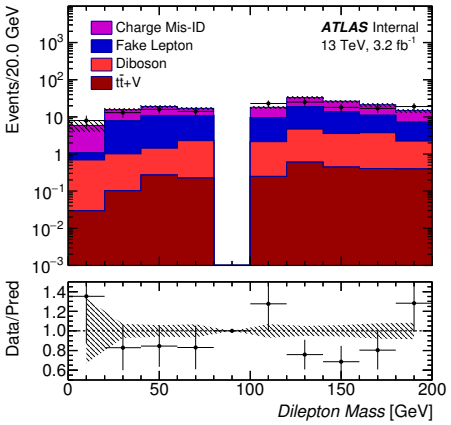
(b) Jet Multiplicity



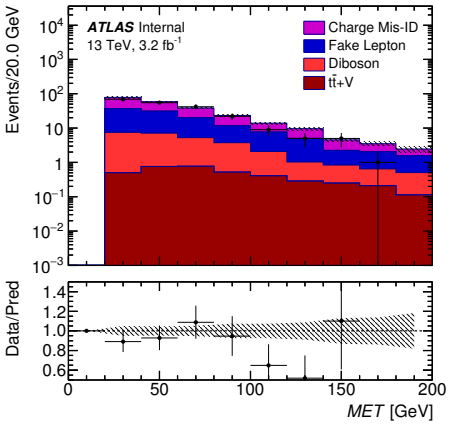
(c) Leading electron p_{T}



(d) Leading electron η



(e) Di-Electron Mass

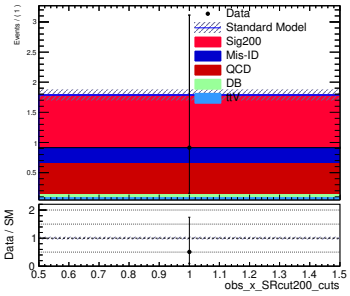


(f) $E_{\mathrm{T}}^{\mathrm{miss}}$

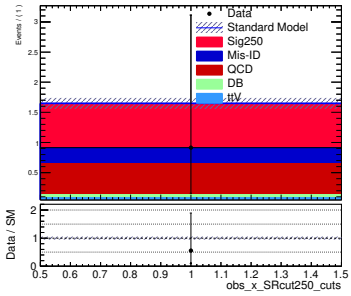
Figure B.2: Same region as above, but with the same-sign normalized inclusive fake contribution.

APPENDIX C. CUT-BASED ANALYSIS

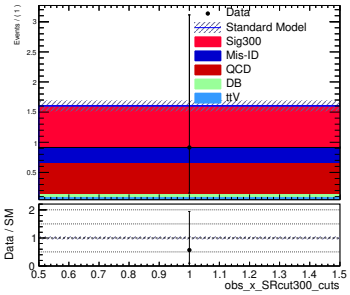
Figure C.1 shows the bins used for the final limit setting for the cut-based analysis.



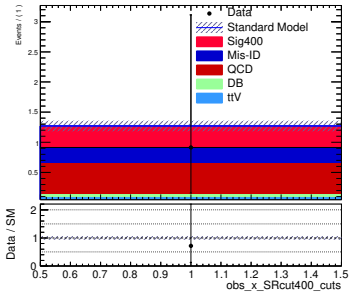
(a) 200 GeVmass point



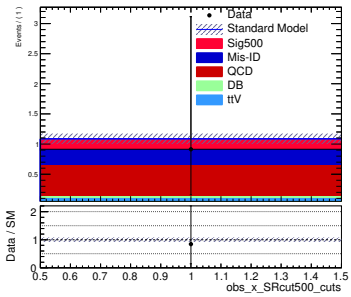
(b) 250 GeVmass point



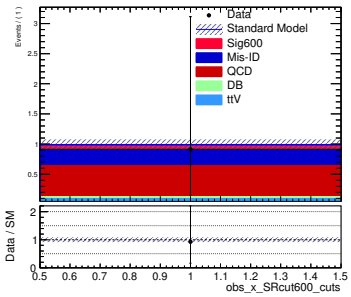
(c) 300 GeVmass point



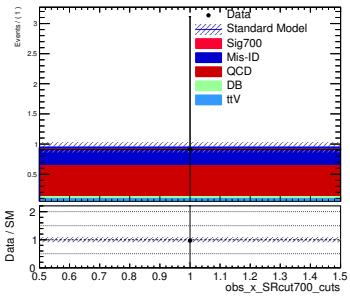
(d) 400 GeVmass point



(e) 500 GeVmass point



(f) 600 GeVmass point



(g) 700 GeVmass point

Figure C.1: Selected events from the cut optimized selection for each of the mass points.